Production and Perception of Laryngeal Constriction in the Early Vocalizations of Bai and English Infants

by

Allison Benner
B.A., Mount Allison University, 1985
Diploma in Applied Linguistics, University of Victoria, 1996

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

in the Department of Linguistics

# SUPERVISORY COMMITTEE

Production and Perception of Laryngeal Constriction in the Early Vocalizations of Bai and English Infants

by

Allison Benner
B.A, Mount Allison University, 1985
Diploma in Applied Linguistics, 1996

**Supervisory Committee**

Dr. John H. Esling, Supervisor
(Department of Linguistics)

Dr. Sonya Bird, Departmental Member
(Department of Linguistics)

Dr. Suzanne Urbanczyk, Departmental Member
(Department of Linguistics)

Dr. Emmanuel Hérique, Outside Member
(Department of French)

# ABSTRACT

**Supervisory Committee**
Dr. John Esling, Supervisor
(Department of Linguistics)

Dr. Sonya Bird, Departmental Member
(Department of Linguistics)

Dr. Suzanne Urbanczyk, Departmental Member
(Department of Linguistics)

Dr. Emmanuel Hérique, Outside Member
(Department of French)

This study examines the production and perception of laryngeal constriction in the early vocalizations of Bai and English infants. The first part of the study documents the development of laryngeal voice quality features in the non-syllabic and syllabic utterances of Bai and English infants. The second part of the study focuses on the perception of laryngeal constriction in infant vocalizations by adult Bai and English listeners. The study is grounded in Esling's (2005) model of the vocal tract, which characterizes the laryngeal vocal tract as a separate articulator, distinct from the oral vocal tract.

The study of Bai and English infants' production identifies universal and language-specific patterns in infants' development of laryngeal constriction. In the first months of life, most sounds produced by Bai and English infants are constricted. As the year progresses, all infants explore degrees of constriction in dynamic utterances that feature alternations between constricted and unconstricted laryngeal voice quality settings. As well, throughout the year, infants produce an increasing proportion of

unconstricted vocalizations. By the end of the first year, when infants have developed increasing control of the laryngeal and oral vocal tracts, they produce syllabic utterances that begin to reflect the use of laryngeal voice quality features in their ambient language. English syllabic utterances are mostly unconstricted, mirroring the prevalence of unconstricted settings in the target language. By contrast, Bai syllabic utterances are mostly constricted or dynamic, reflecting the use of laryngeal voice quality in Bai, a register tone language that employs laryngeal voice quality features distinctively at the syllabic level.

The second part of the study highlights universal and language-particular patterns in Bai and English adults' perception of laryngeal voice quality in infants' utterances. In evaluating the importance of a range of infant sounds in learning the target language (Bai or English), adults from both language groups assign lower ratings to infant utterances that occur earlier in development, such as constricted non-syllabic utterances, and higher ratings to sounds that occur later, such as syllabic utterances with rapidly articulated syllables. Bai and English adults' perceptions also reflect some language-specific patterns that correspond to language-particular characteristics identified in infants' use of laryngeal voice quality in syllabic and non-syllabic utterances. These correspondences suggest that adults are attuned to laryngeal voice quality in infants, and that, in turn, infants become attuned to the use of laryngeal voice quality features in their ambient language early in development.

The production study demonstrates the fruitfulness of Esling's (2005) model of the vocal tract in revealing previously undocumented patterns in the development of laryngeal constriction in the first year of life and in highlighting the importance of emergent laryngeal control as a stimulator of phonetic development. The perception study shows that adults whose native languages differ markedly in their use of laryngeal constriction can systematically evaluate laryngeal voice quality features in the full range of non-distress vocalizations produced by infants in the first year of life.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ACKNOWLEDGMENTS

I take a perhaps dubious pride and pleasure in noting that I have spent nearly one-quarter of my life as a graduate student in the Department of Linguistics. While I cannot possibly do justice to the generations of students, faculty, and staff who have contributed to my experience, please accept my thanks.

I would like in particular to thank my supervisor, Dr. John Esling, for his support, guidance, and encouragement in all my endeavours, including this dissertation. Thanks, too, for the wonderful opportunities to see a little more of the world while getting my work done. I would also like to thank Dr. Suzanne Urbanczyk, Dr. Sonya Bird, and Dr. Emmanuel Hérique for agreeing to serve on my committee, and for all your efforts to support me, particularly in the final throes. I would also like to thank Dr. Janet Werker for agreeing to be my External Examiner.

Thank you to the many friends in the department whose support and companionship has been so important in the last couple of years. Janet, Izabelle, Jun, Scott, and Thomas, you have played particularly important roles in keeping my spirits up. A particular thanks must go to Qian, without whom I could not have completed my research in Yunnan, China last year. I would also like to thank my many friends outside the department who have supported and inspired me throughout my graduate studies and in the rest of my life. Michelle, Rhonda, Robin, Debbie, and David all deserve special mention for “being there” for me in more ways than I can say here.

Finally, I would like to thank several members of my inner circle for their support. Loucas, thank you for being a good father to Anna during these challenging times, and for your help in producing an illustration for this dissertation on short notice. Thank you, Mom, for showing me how to cheerfully defy the odds in the face of life’s challenges—your example will inspire me in good times and bad for the rest of my life. Dad, the memory of your warmth and love of people reminds me of what motivates and sustains me in most of my work, quite aside from its inherent interest. Leah, thanks for being available anytime to hash over the events of the hours, days, and weeks. Jen, thanks for both the conversations—and the reflective silences—over the last few years. They have served a welcome reminder of the need to keep taking deep breaths. And finally, a heartfelt thanks to Peter for the gentle and steady encouragement to take another step in my work, especially at times when I could no longer see the point. May all our projects end so well.

It’s been fun, I think. Thanks to everyone for that.

# DEDICATION

I would like to dedicate this dissertation to my daughter, Anna, who has inspired me in countless ways to keep striving to improve myself (I know ... I’m not done yet).

# Chapter 1

# INTRODUCTION

### *1.1 Purpose of the Study*

The purpose of this study is to investigate the relationship between infants' production of voice quality features in the first year of life and adult perceptions of voice quality features in infant vocalizations. Based on perceptual studies, it is widely recognized that over the course of speech development, infants attune themselves to their ambient language, beginning their lives with the capacity to recognize all segmental distinctions produced in human language, and eventually losing this capacity in the second half of the first year in favour of recognizing native language contrasts (Kuhl et al., 1992, 2006, Werker & Curtin, 2005; Werker & Tees, 1992, 1999). However, a much smaller body of research has focused on the question of whether, at what point, and in which respects, infants' utterances begin to reflect the influence of the ambient language, though some research suggests that prosodic features are among the first phonetic features to exhibit language-specific characteristics (Boysson-Bardies, et al., 1984; Hallé & Vihman, 1991; Nathani et al., 2003; Zhu & Dodd, 2000). To date, however, no phonetically grounded, cross-linguistic study has focused on the development of voice quality features in the first year of life. This study addresses this gap in the literature.

Throughout the first year, adults serve as infants' primary caregivers and important models of the ambient language(s) infants are acquiring. As caregivers of infants, adults interact vocally with infants on a daily basis, possibly encouraging vocal features they perceive to be important in the surrounding language and culture, and discouraging the proliferation of other features that may play a less important role in the language or that may be regarded unfavourably within the culture (Locke, 2006; Watson, 1972, 1979, 1985). Yet, little is known about how adults perceive the sounds that infants make. While some studies have focused on adults' perceptions of babbling (Boysson-Bardies et al., 1984; Bloom & Lo, 1990; Bloom et al., 1993; Goldman, 2001; Oller et al., 2001) and while many studies have focused on adults' perceptions of crying (Bisping et al., 1990; Frodi & Lamb, 1980; Frodi & Senchak, 1990; Murray, 1985; Papousek & Von Hofacker, 1998; Zeskind, 1987; Zeskind & Shingler, 1991), few studies have examined adults' perceptions of the full range of sounds produced by infants in the first year. While studies of crying highlight the role of laryngeal constriction in shaping adults' perceptions (see, for example, Zeskind & Shingler, 1991), no studies have focused on specifically on adults' perception of laryngeal voice quality features in infants' non-distress vocalizations. It is of particular interest to know how such perceptions may differ according to the use of laryngeal voice quality features in the ambient language, and whether these perceptions bear a systematic relationship to the development of voice quality features in infants' utterances during the first year. The present study provides a foundation for research in this area.

This study is comprised of two parts. The first part of the study examines the production of voice quality features in infant vocalizations among infants from two languages that differ markedly in their use of voice quality features: English, a Germanic language that uses voice quality non-contrastively as a means of paralinguistic expression

(Esling, 1994, 2000; Laver, 1980, 1994); and Bai, a Tibeto-Burman language that employs voice quality distinctively at the syllabic level as part of its register tone system (Esling & Edmondson, 2002). The second part of the study is a qualitative and quantitative analysis of the perception of selected voice quality features in a range of infant vocalizations by English and Bai adult listeners.

This study addresses a number of gaps in the literature. First, building on the foundation established by other infant speech researchers who have examined speech development from birth (Oller, 1978, 1980; 2000; Roug et al., 1989; Stark et al., 1975, 1993; Nathani et al., 2006), this study provides a principled phonetic approach to the classification of infant vocalizations, drawing on Esling's (1996, 1998, 1999a, 1999b, 2005; see also Esling & Harris, 2003) model of the vocal tract and associated research into the use of laryngeal constriction in languages of the world (Carlson et al., 2004; Edmondson & Esling, 2006; Edmondson et al., 2005; Esling & Edmondson, 2002). Second, by focusing specifically on laryngeal voice quality in infant vocalizations, this study addresses a neglected and poorly understood aspect of speech and language development in infancy. Third, this study examines the use of voice quality features among infants learning English, a much-studied language, with the use of these same features among infants learning Bai, a much less-studied language, extending our understanding of universal and language-particular aspects of speech and language acquisition. Fourth, this study explores how adults from two different language backgrounds perceive voice quality in a range of infant vocalizations that occur in the first year of life. Finally, this study examines whether there is a systematic relationship between infants' production of voice quality features and adults' perceptions of those features, particularly towards the end of the first year of life, when language-particular patterns may begin to emerge.

## 1.2 Research Questions

### *1.2.1 Study 1: Infant Production*

The study of English and Bai infants' production of voice quality features explores the following questions:

- **Question 1:** How does the use of laryngeal voice quality features, particularly the use of laryngeal constriction, change over the first year of life?
- **Question 2:** Does the incidence of laryngeal constriction differ according to the type of utterance that infants produce? In particular, does the use of laryngeal constriction differ in syllabic and non-syllabic utterances in general and over time?
- **Question 3:** What is the relationship between the development of laryngeal control and the emergence of control within the oral vocal tract?
- **Question 4:** Are there universal patterns in the development and use of laryngeal voice quality features in the first year of life?
- **Question 5:** Does infants' use of laryngeal voice quality features differ according to the use of these features in the infants' ambient language? If so, do these differences affect other features of infant speech development, such as the development of babbling and the development of control over the oral vocal tract?

### *1.2.2 Study 2: Adult Perception*

The main questions of the perception component of the present study are:

- **<u>Question 1</u>:** Can phonetically untrained adults systematically evaluate infant vocalizations outside of specific relationships with infants they know?
- **<u>Question 2</u>:** Are phonetically untrained adults sensitive to laryngeal voice quality features in evaluating infant vocalizations?
- **<u>Question 3</u>:** Are there universal patterns in phonetically untrained adults' perceptions of laryngeal voice quality features in infant vocalizations?
- **<u>Question 4</u>:** What is the relationship between adults' ratings of infant vocalizations and universal processes of infant phonetic development in production? Do higher ratings correlate with features that infants control later in the process of phonetic development?
- **<u>Question 5</u>:** Do language-specific factors affect phonetically untrained adults' perceptions of laryngeal constriction in infant vocalizations? If language-specific factors exist, do these factors relate systematically to language-specific patterns in the use of laryngeal constriction in infants' production?

## 1.3 Limitations of the Study

This study is primarily exploratory in nature. The first part of the study, which addresses infants' production of voice quality features, is based on the utterances produced by four infants from each language studied (Bai and English). The findings of the study will need to be extended to larger groups of infants within these language groups, and to infants from other language groups that differ in their use of laryngeal voice quality features.

Similarly, the perceptual component of the present study, which addresses adult perceptions of infant vocalizations, is based on the perceptions of 40 adults from each language group. Further research is necessary with larger groups of adults from the language groups to control for factors not addressed in this study, such as education and socio-economic status. Also, because the perceptual component of the study used natural infant vocalizations, it was difficult to control for the many segmental and suprasegmental features that might have affected adults' perceptions of the sounds, limiting the conclusions that can be drawn from the research. However, this study helps to identify factors that may affect adult perceptions of infant vocalizations and that can be incorporated into future research using natural infant speech stimuli.

## 1.4 Outline of Dissertation

This chapter provides a general overview of the production and perception components of the present study, and outlines the main research questions that have guided each part of the study. The second chapter of this dissertation provides a review of the literature. The third chapter of this document outlines the methodology employed in the present study. In the fourth chapter, the results of the analysis of the study of infant vocalizations are presented and discussed. The fifth chapter describes and discusses the results of the perception study, including the quantitative and qualitative portions of the investigation. Chapter six features a discussion of the relationship between the study of infant production and adult perception of infant vocalizations. Chapter 7 provides a summary of the main findings of the present study, and outlines directions for future research.

# Chapter 2

# LITERATURE REVIEW

In this chapter, I review the literature on infant speech development relating to production and perception. In Section 2.1, I review research on infant speech production in the first year of life. In Section 2.2, I review the literature on infant speech perception and adults' perception of infant vocalizations. In Section 2.3, I outline the phonetic research that provides the theoretical foundation for the present study. Finally, in Section 2.4, I conclude this chapter by describing how the literature has shaped the focus of the present study, and what gaps in existing knowledge this study addresses.

## 2.1 Production

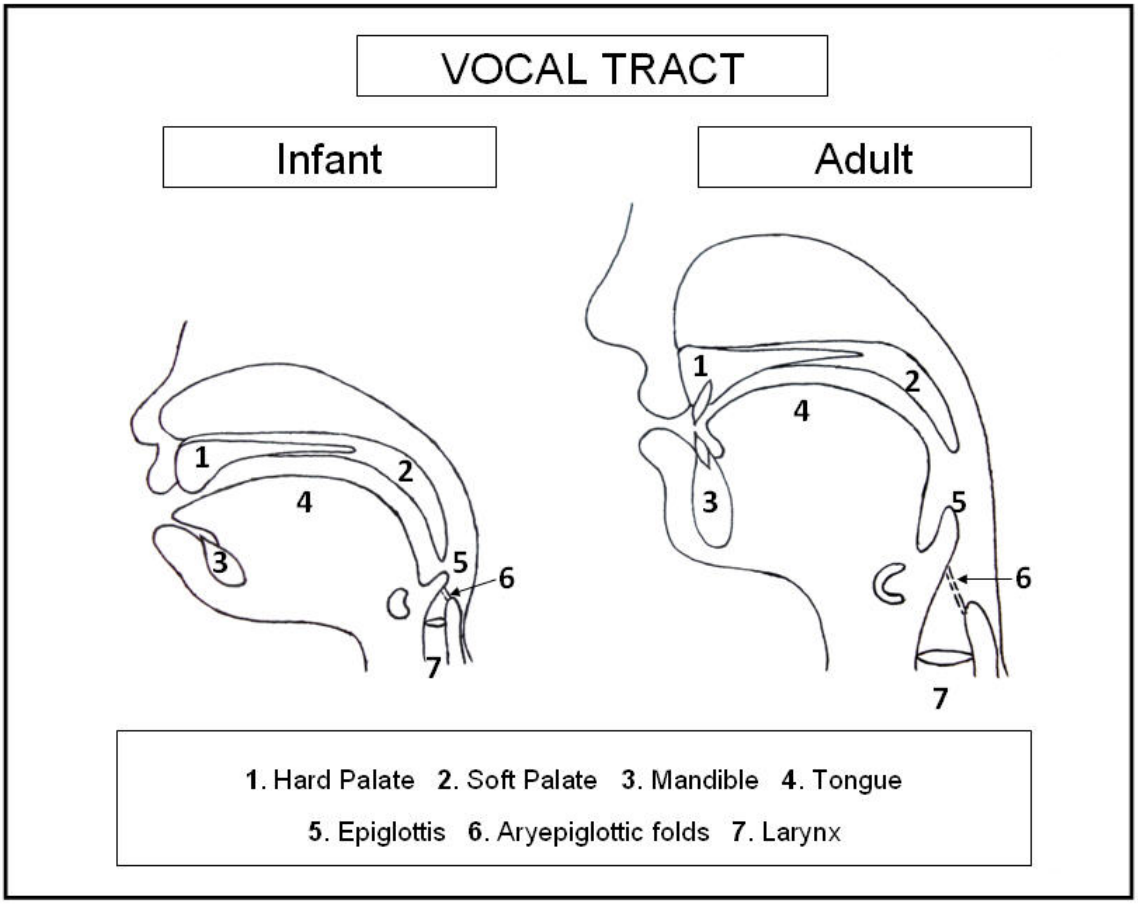
### *2.1.1 Infant Vocal Tract*

The disposition of the infant larynx throughout the first year of life, in combination with psychological factors to be discussed in Section 2.1.2, directly impacts the nature and development of infant vocalizations in the first year of life.

Most information on the disposition of the infant larynx stems from the medical literature (Crelin, 1973; Eckel et al., 1999; Fried et al., 1982; Kent & Vorperian, 1995; Sasaki et al., 1977; Vorperian et al., 1999), with some additional research originating from researchers with a particular interest in the relationship between infant vocal physiology and the evolution of language (Fitch & Giedd, 1999; Lieberman et al., 2001). According to this body of research, the infant vocal tract differs from the adult vocal tract in five main respects. First, in constrast to the right-angled posture of the adult oral cavity relative to the pharynx, the infant oro-pharyngeal channel is sloped. Second, the tongue is in a retracted position, by virtue of its slope and predisposition to engage with supraglottic structures. Third, the vocal folds are short and the musculature that controls them is undeveloped. Fourth, compared to adults, the larynx is in a higher and more constricted position. Finally, the epiglottis rests against the soft palate, maintaining a respiratory passageway from the larynx through the posterior nares. Starting in the third month of life, the infant vocal tract undergoes significant growth and restructuring, increasing infants' vocal capacities. The larynx, epiglottis, and hyoid bone begin to descend, lengthening the vocal tract and changing the angle of the oral cavity relative to the pharyngeal cavity. This restructuring continues throughout infancy, affording infants the capacity to produce a wide range of pitch, intensity, and constriction at the laryngeal level.

Figure 2-1 below illustrates the differences between the infant (newborn) and adult vocal tracts. This illustration, based on descriptions and illustrations provided in Kent and Murray (1982) and Kent and Vorperian (1995), shows the location of major anatomical structures in the vocal tract, including the aryepiglottic folds. The illustration also depicts the relationship between the laryngeal and oesophageal passages: when the laryngeal passage is open, as shown in Figure 2-1, the oesophageal passage is closed, and vice versa.

**Figure 2-1. Infant versus Adult Vocal Tract**




The anatomical features of the infant vocal tract make it more similar to the mature larynges of other primates than to the mature larynges of mature members of their own species. Given the comparatively limited vocal abilities of other primates compared to humans—and in particular, given the absence of language among the other primates—many researchers are led to the seemingly natural conclusion that the infant vocal tract is “underdeveloped” or “primitive” compared to the adult vocal tract (Kent, 1981; Thelen, 1991).

While it is true that in the first months of life, infants are unable to produce the range of oral sounds that are featured in language, from the time they are born, infants are physiologically disposed to produce a range of stricture-based and prosodic sounds that involve laryngeal constriction, many of which are employed in human languages (Carlson et al., 2004; Catford, 1977; Edmondson & Esling, 2006; Edmondson et al., 2005; Esling & Edmondson, 2002; Ladefoged & Maddieson, 1996). Moreover, as some researchers are now discovering, the right-angled, lowered larynx is not limited to humans (Fitch & Reby, 2001), and is, thus, not necessarily the primary determinant in the evolution of speech and language. While, as we will see, the disposition of the infant larynx certainly constrains the range of sounds that infants can make, particularly in the first months of life, the *way* that infants use their developing vocal capacities may illustrate properties that are at least as relevant to human speech and language development as any particular physiological setting. To better understand these features, we now briefly review the psychological research on infant learning in social contexts.

### *2.1.2 The Role of Vocal Imitation in Social Learning*

Research in the past 20 years has dramatically changed our understanding of infants' cognitive capacities, revealing that many features that are traditionally thought of as sophisticated, late-developing abilities, such as the capacity to engage in collaborative, goal-directed behaviour, begin to develop in the first year of life (Gergely & Csibra, 2004; Tomasello et al., 2005). Mutual attunement of infants and adults to one another's behaviour is a powerful stimulus to the development of goal-oriented behaviour and the formation of shared intentions between infants and their caregivers. In this interactive process, infants need to develop a sense of identification with people in their environment (Hauf & Prinz, 2005). Imitation is one way that infants demonstrate, develop, and explore the relationship between what other people do, and what they can do, at any given point in development (Gergely & Csibra, 2005).

Following up on an earlier line of research about infants' abilities to imitate facial expressions from the first days of life (Meltzoff & Moore, 1977, 1983, 1989), Meltzoff and Moore (2002) examined six-week-old infants' capacities to imitate oral facial expressions (e.g., mouth opening, tongue protrusion); to retain the memory of other people's facial expressions; and to produce increasingly close imitations of previously observed and produced gestures upon repeated exposure to the same person. In line with their previous studies, researchers found that young infants imitate facial expressions, and that when they see the same person producing the same gesture, they gradually modify their imitations to produce more accurate matches over time. The authors hypothesize that infants engage in motor imitation as an early way of understanding and communicating with other people.

Some evidence shows that infants imitate the vocal output of other people quite early in development. Kuhl and Meltzoff (1996) examined infants' vocalizations in response to adults' productions of the vowels /i/, /a/, and /u/ at 12, 16, and 20 weeks of age. While the authors found developmental changes in infants' capacity to produce these vowels distinctly over time, they also found that infants responded with vocalizations that perceptually matched the vowels presented to them, even at 12 weeks of age.

While infants must gradually attune themselves to people in their surroundings, caregivers also attune themselves to infants. The capacity of infants and caregivers to become mutually attuned to one another increases as infants' vocal capacities expand. For example, Ginsberg & Kilbourne (1988) found that the incidence of interactive exchanges between mothers and infants increases at the end of the third month, around the time that infants acquire greater control over pitch, intensity, and constriction. These exchanges, while not specifically linguistic in nature, are examples of social sound-making that may be precursors to language, both in the course of infant development and in the evolution of language within our species (Locke, 1998).

Caregivers' attunement and response to infants' vocalizations is one means by which infants may come to understand their utterances as intentional (Meadows et al., 2000). Getting a consistent response from a caregiver in response to a certain type of vocalization, and hearing the parent muse out loud about the ascribed meaning of that vocalization, creates a pairing of sound and meaning that is a precursor to the development of linguistic meaning. At the same time, caregivers choose, over time, to selectively reinforce certain vocal behaviours over others, in keeping with their knowledge and understanding of infants' evolving vocal and communicative abilities (Watson, 1972, 1979, 1985). As proposed by Locke (2006), while parents may respond to a newborn's harsh cry, and while such cries may be uniquely adapted to attract caregivers' attention (Fitch et al., 2002; Furlow, 1997; Lieberman et al., 1971; Lummaa et al., 1998), infants who do not eventually produce more cooing than crying, and later, more babbling than cooing, may not receive the same attention from parents as infants who produce these sounds within the typical time frame, and may be at increased risk of neglect, abuse, and infanticide.

### *2.1.3 Exploratory Vocal Play*

As highlighted by Locke (2006), cited above, infants have a powerful incentive to engage in progressively more complex acts of vocal exploration, and caregivers have a powerful incentive to encourage them to do so. It is important to note that while imitation may be a powerful means of social learning, vocal exploration is not imitative. As noted by previous infant speech researchers (Oller, 2000), infants most typically engage in vocal exploration when they are relaxed and solitary, and are disposed to explore and build upon whatever vocal resources are at their command at any given point of development.

Research shows that in the first months of life, crying is the most frequent type of vocalization produced by infants, and that crying exhibits a high degree of laryngeal constriction, though it is not generally named as such. Some of the most detailed acoustic studies of infant vocalizations focus on crying. Earlier research into infant crying focused on identifying markers of pathology, and highlighted the presence of “abnormal,” “chaotic” patterns such as sudden changes in pitch and/or subharmonics as indicators of poor health status (Barr, 1990, 1998; Gilbert & Robb, 1996; Goberman & Robb, 1999; Grauel et al., 1990; Thóden et al., 1985; Zeskind & Barr, 1997). However, as highlighted by Buder et al. (2008), more recent studies (see, for example, Robb, 2003) suggest that very harsh and/or high-pitched cries once thought to be indicators of pathology are increasingly understood to be part of the vocal repertoires of normally developing infants. However, even in the first months of life, the crying of typically developing infants grows in melodic complexity and changes in resonance, suggesting that even in this most “undeveloped” of vocalizations, which involves a reflexive use of pitch and laryngeal constriction, infants explore and develop their vocal capacities (Wermke et al., 2002). While infants' earliest sounds may well be reflexive, infants' innate motivation to learn may prompt them to use whatever vocal abilities they possess as a launching point for exploration.

Oller (1978, 1980, 2000, 2004), one of the most influential infant speech researchers of the past 30 years, has emphasized the role of "vocal play" or "vocal entertainment" in speech development. Between the ages of three and seven months, infants are often seen to engage in systematic alternations between new sounds, including different vowels, different pitches, and different phonation types. Oller (2000, 2004) characterizes such vocal play as the infant's first exploration in contrastive features, and thus, as the precursor to the development of linguistic categories. Other researchers have also noted increased exploratory vocalization during these months (Koopmans-van Beinum & van der Stelt, 1986; Roug et al., 1989; Stark, 1980; Thelen, 1991). In a detailed study of the vocalizations of one English infant, Bettany (2004) highlighted the role of alternations in laryngeal voice quality in developing laryngeal control in the first six months of life, observing that the number and length of such alternations increase in the fourth month of life, coinciding with a period of increased growth and development in the larynx.

### *2.1.4 The Emergence of Syllabic Vocalizations*

Some time after, or overlapping with, the period of vocal play, infants begin to produce canonical babbling sequences, utterances that feature repetitive consonant-vowel (CV) syllables. In the initial stages of babbling, sometimes referred to as "marginal" babbling (Oller, 1980), the consonants in babbling sequences do not display the timing features of adult CV(C) syllables, with stops showing longer periods of silence and longer transitions to vowels. With practice, however, CV(C) sequences begin to acquire the timing features seen in adult utterances, allowing infants to produce babbling sequences with rapidly articulated syllables. According to Koopmans-van Beinum and van der Stelt (1986), babbling begins to appear in infant vocalizations by the seventh month of life in 90% of infants. According to Oller et. al. (1998, 1999), the failure to produce babbling by this time is often predictive of later speech and language disorders.

MacNeilage's "frames, then content" theory of speech development (1998) has been profoundly influential in shaping infant speech researchers' understanding of the origins of babbling. MacNeilage and his colleagues (MacNeilage et al., 2000, 2001) have explored the notion that babbling originates from the jaw cyclicities involved in chewing, and thus, evolves from an adaptation of a human non-speech behaviour. The "content" of early babbling—the specific vowels and consonants—is initially dominated by the "frame" provided by jaw movement, making bilabial babbling sequences produced with low central vowels among the first to appear in the productions of most infants. Over time, infants begin to produce alveolar and velar babbling sequences, the latter often produced with back vowels and the former with front vowels, in keeping with the notion of "frame dominance." With practice, infants gain freedom in producing different combinations of consonants and vowels, allowing for the integration of language-specific "content" in syllabic utterances, including language-specific prosody (Davis et al., 2000).

### *2.1.5 Prosodic Development*

Of all the utterances infants produce in the first year of life, babbling is considered to be the most speech-like and, therefore, the most likely among infant vocalizations to reflect the influence of the ambient language. Many researchers have examined the babbling of infants from different language backgrounds, with a view to identifying language-specific features. Consistent with MacNeilage's (1998) perspective, the segments in infant babbling tend to show a remarkable consistency in early babbling, but the prosody may feature language-specific characteristics (Boysson-Bardies et al., 1984; Hallé & Vihman, 1991; Nathani et al., 2003), and the control of some aspects of laryngeal control, such as prevoicing and/or voice onset time, may begin to show in the ways that infants produce some consonants in babbling sequences (see, for example, Whalen et al., 2007). Overall, however, while research suggests that infants actively explore linguistically relevant aspects of prosody in the first year (Delack & Fowlow, 1978; Hsu et al., 2000; Sheppard & Lane, 1968), sometimes producing utterances that resemble their ambient language (Hallé & Vihman, 1991), they do not begin to consistently combine these features in language-specific ways until at least the second year of life (Snow & Balog, 2002).

Analysis of the prosodic features in babbling tends to focus on the role of pitch, length, and loudness. One obvious, but less studied, area of suprasegmental development is the development of tone. Most studies of tone acquisition are based on studies of various dialects of Chinese (see, for example, Clumeck, 1977, 1980; Li & Thompson, 1977; and Ota, 2003). This body of research consistently demonstrates that while infants play extensively with pitch in the first year of life, they do not acquire specific lexical tones until well into the second year of life. However, this research sheds light on the trade-offs involved in learning a language in which prosodic features carry a heavy load. In an analysis of speech errors produced by Chinese children, starting in the second year of life, Zhu and Dodd (2000) found that tones were acquired first, followed by syllable-final consonants and vowels, and then by syllable-initial consonants—a pattern of errors that would be unusual among children learning English, who would be least likely to make errors in syllable-initial consonants at this age. The authors suggest that contrary to the traditional position articulated by Jakobson (1941/1968), the *salience* of a given feature within the language may influence the order of acquisition independent of its *markedness* in languages generally.

### *2.1.6 Acquisition of Voice Quality*

The acquisition of voice quality in infancy and childhood is seldom considered (though see Foulkes et al., 2001, which examines this issue in older children), except as reflected in a bias towards studying sounds produced with modal voice. In the past 35 years, infant speech researchers have paid increasing attention to the early vocalizations of infants as precursors to communication and language development. In earlier research, studies focused on identifying the most “speech-like” early vocalizations, so many researchers focused only on earlier utterances produced with “normal” or modal phonation (Koopmans-van Beinum & van der Stelt, 1986; Oller, 1980; 2000). Oller (2000) considers modal phonation as a first step in the development of control over syllable production in babbling. Buder et al. (2008) note that “caregivers, researchers, and others primarily interested in tracking incipient language understandably attend to productions spoken with modal voice, as being indicative of emerging linguistic control, while treating squealy or growly voices as pertaining to more paralinguistic communication indicating emotion, attitude, or overall fitness” (p. 553).

However, it is increasingly recognized that these “squealy” or “growly” sounds, which are non-modal sounds produced with laryngeal constriction, represent a large proportion of the sounds produced in early infancy; and that they may play an important role in the evolution of infants’ communicative resources (McCune et al., 1996) and in their phonetic development (Bettany, 2004; Esling et al., 2004a, 2004b, 2004c). Moreover, some languages use these features contrastively (Esling & Edmondson, 2002). For adults who speak such languages, modal voice may not be the hallmark of “typical adult speech” (Buder et al., 2008) or of canonical syllable production. In addition, consistent with the perspective voiced by Zhu and Dodd (2000), infants who are learning such languages may find these sounds particularly salient, and caregivers may encourage their production. To better explore this possibility, we now turn to the research on infant speech perception and on adults’ perceptions of infant vocalizations.

## 2.2 Perception

### *2.2.1 Infant Speech Perception*

The last 30 years have seen a dramatic increase in our understanding of the perceptual capacities of infants in the first year of life. As reviewed above, while infants’ patterns of production may reflect some properties of the ambient language, the evidence for the effect of the target language is stronger in studies of infant speech perception. Several reviews of infant speech perception highlight the mounting evidence that infants become increasingly attuned to the phonetic features of their native language (Werker & Tees, 1984, 1992, 1999). Research shows that in the first six months of life, infants are capable of recognizing the distinctions that are made in virtually all human languages. However, by the second half of the first year, infants lose this sensitivity, in favour of increasing attunement to the contrasts that function in lexical contrasts in their target language (Kuhl, 2000; Kuhl et al., 1992, 2006). This effect is evident both in infants' capacity to distinguish vowels (Polka & Werker, 1994) and consonants (Eimas, 1974; Eimas et al., 1971; Eimas & Miller, 1980). A similar effect has been shown in infant tone perception, though the loss of sensitivity to non-native tonal contrasts may occur somewhat later than is the case for segmental distinctions (Mattock & Burnham, 2006; Mattock et al., 2008). Indeed, as demonstrated by Kuhl et al. (2005), infants who do not lose their capacity to distinguish non-native sounds in the second half of the first year are slower to develop language abilities later in infancy.

A further body of research demonstrates infants' sensitivity to prosodic cues in their target language and the importance of prosodic representations in early lexical representations. Jusczyk et al. (1993) and Jusczyk & Kemler Nelson (1996) have shown that English infants prefer sounds that exemplify the predominant stress patterns employed in English words. This finding is consistent with that reported by Vihman et al. (2004), in their study of the role of accentual patterns in word recognition among English- and French-learning infants. These tendencies, combined with infants' demonstrated preference for lexical words early in infancy (Shi & Werker, 2001), suggest that infants are highly sensitive to segmental and suprasegmental cues that play a distinctive role in their language.

Based on these trends in the research, it is reasonable to predict that infants who are learning a language where laryngeal voice quality is distinctive will be highly attentive to pitch-related variations in voice quality in their ambient language and that adults who speak such languages may well emphasize these distinctions in child-directed speech and encourage their proliferation in infants' vocalizations. However, these comments are purely speculative, as there is almost no research on infant speech production in such languages (though see Benner et al., 2007, and Esling et al., 2006) and no research at all on adult perception of infant vocalizations in these linguistic contexts.

### *2.2.2 Adult Perception of Infant Vocalizations*

Very little is known about how adults perceive the full range of vocalizations produced by infants in the first year of life. Most research on adult perceptions of infant vocalizations focuses on their responses to babbling or to crying. The former body of research tends to exclude babbling produced with laryngeal constriction (see, for example, Boysson-Bardies et al., 1984), while the latter studies focus almost exclusively on sounds produced with laryngeal constriction, though only of the variety produced when infants are distressed.

Oller (2001) has noted that parents with a wide range of educational and socio-economic backgrounds easily recognize the onset of canonical babbling in their infants. Moreover, adults have been shown to prefer syllabic vocalizations to other infant sounds, perceiving infants who produce a high rate of syllables per vocalization as more pleasant, friendly, and likeable than infants who produce fewer syllables or who produce primarily vocalic utterances (Bloom & Lo, 1990; Bloom et al., 1993). This preference may have evolutionary roots in other primate caregivers' preferential responses to reduplicative vocal patterns produced by their young (Elowson et al., 1998a, 1998b). Locke (2006) cites a particularly revealing example of the role of babbling in facilitating parental engagement with infants. He cites one American woman's description of the turning point in her decision to adopt an infant:

> She was not talking at all when we met her," the mother said in a letter, "except for making primitive grunting sounds. She made no eye contact and appeared extremely withdrawn and afraid of her surroundings. When we put her down on the floor to play she would only crawl to [a] corner and sit facing [the] wall or just rock and roll on [her] back. One time I handed her back to [the] caregiver and just before leaving for the day, I said bye-bye and she yelled out loud and clear, 'MA MA!' Even the caregivers were surprised and stated they had never heard her speak before that. My husband and I were so taken by this that then and there we decided to do everything possible to bring her home and help her. (Locke, 2006: 162-163).

Adults' preference for syllabic vocalizations may be contrasted with their preferences for "cooing," vocalic utterances produced without laryngeal constriction, and their aversion to crying, vocalic utterances produced with laryngeal constriction. In the first months of life, infants cry frequently. Towards the end of the second month of life, however, the frequency of crying tends to decrease in favour of the increased production of cooing (Hopkins & von Wulffften Palthe, 1987), which, in turn, increases caregivers' engagement in vocalic turn-taking with their infants (Ginsburg & Kilbourne, 1988; Watson, 1972). Research shows that infants who do not begin to produce these nasalized, unconstricted vocalic sounds at the expected times in development may receive less care than their cooing counterparts (Murray, 1985; Papousek & von Hofacker, 1998; Papousek et al., 2001). Persistent infant crying increases the heart rate among men (Frodi & Lamb, 1980; Zeskind, 1987) and increases the production of testosterone among men and women alike (Fleming et al., 2002), possibly increasing the likelihood of aggression (Zeskind, 1987).

Laryngeal constriction may play a specific role in stimulating these negative responses to infant crying. As noted earlier in this review, infants who cry more frequently than others are at increased risk of abuse, neglect, and infanticide (Frodi, 1985; Frodi & Lamb, 1980; Frodi & Senchak, 1990). Adults appear to respond especially negatively to infant cries that feature an aperiodic, “dysphonated” quality (Zeskind & Barr, 1997) which is characteristic of a strong degree of laryngeal constriction. Adults who have been charged with child abuse have sometimes cited the “grating sound of the cry” as the precursor to the abuse (Frodi 1985, cited in Locke, 2006).

It seems possible that, at least among infants who are learning English, increased production of unconstricted vocalizations in infancy is associated with reductions in crying, increased vocal interactions with caregivers, and the development of more controlled syllabic utterances. However, it is unclear whether the relationship between laryngeal constriction and speech development, including adults’ perceptions of speech development, is similar in environments where the language itself features laryngeal constriction. Moreover, even among English infants, the role of laryngeal voice quality variations in speech development is unclear. While English adults’ responses to laryngeal constriction in crying may be understandably negative, it is not clear that their responses to similar phonetic features in non-distress vocalizations, whether syllabic or non-syllabic, are negative. Further research is necessary on this issue within a phonetically grounded perspective. For that, we now turn to a discussion of phonetic research into laryngeal constriction.

## 2.3 Phonetic Studies of Laryngeal Constriction

Many researchers have noted the strong presence of laryngeal constriction in the non-distress vocalizations of infants, though this feature has been described inconsistently, using a wide range of impressionistic terms. Bettany (2004) compiled a list of such terms in her study of laryngeal constriction in the early vocalizations of one English-learning infant. She identified a range of semi-phonetic terms used to describe laryngeal constriction, including “hyperphonation” and “aperiodic glottal excitation” (Lieberman et al., 1971), “dysphonation” (Möller & Schönweiler, 1999); “glottal pulses” and “pharyngeal friction” (Stark et al., 1975); and “vocal tremor” (Kent & Murray, 1982; Möller & Schönweiler, 1999).

More typically, however, sounds with laryngeal constriction are described in purely impressionistic terms. These terms include “squealing” (Oller, 1980, 2000); “screaming” (Koopmans-van Beinum & van der Stelt, 1986; Buder et al., 2003); “shrieking” (Koopmans-van Beinum & van der Stelt, 1986); “growling” (Kent & Murray, 1982; Robb et al., 1989; Oller, 2000); “moaning” (Scheiner et al., 2002); “coughing” (Oller, 2000); “grunting” (McCune et al., 1996); “small, throaty sounds” (Stark et al., 1975); and finally, “chaos” (Buder et al., 2003; Fitch et al., 2002). By contrast, unconstricted sounds produced with modal phonation are generally described as “normal” (Buder et al., 2003; Oller, 1980; 2000), despite the fact that they are not the dominant phonatory mode in infancy, among normally developing infants (Bettany, 2004; Benner et al., 2007).

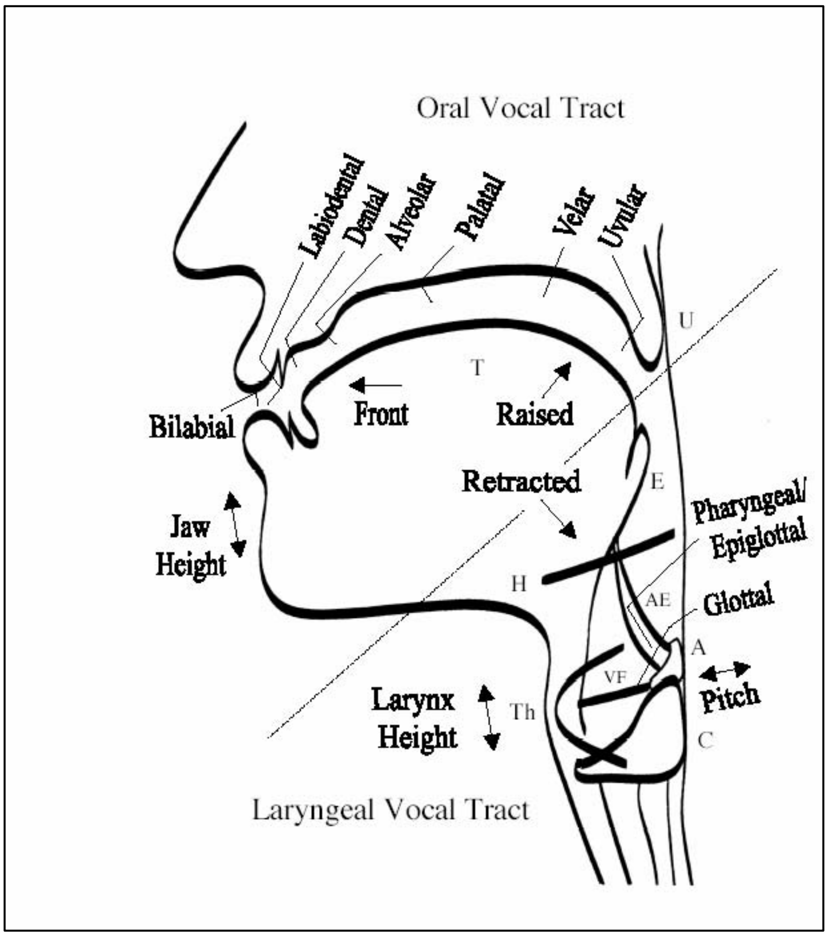
The work of Esling and his colleagues has helped to clarify the physiological mechanisms underlying laryngeal constriction (Esling, 1996, 1998, 1999a, 1999b, 2005) and their use in various languages of the world (Carlson et al., 2004; Edmondson & Esling, 2006; Edmondson et al., 2005; Esling & Edmondson, 2002). Bai, a Tibeto-Burman language spoken in Yunnan, China, provides a particularly striking example of the contrastive use of laryngeal constriction. While English employs laryngeal voice quality features primarily as a means of paralinguistic expression (Esling, 1994, 2000; Laver, 1980, 1994), Bai uses laryngeal constriction throughout its tone system at the syllabic level. As documented by Esling and Edmondson (2002), laryngeal constriction interacts with pitch and nasality to create a tonal paradigm that includes 15 distinctive tones. Table 2-1 below, adapted from Esling and Edmondson (2002), shows the relationship between pitch and laryngeal constriction in the Bai tonal paradigm.

**Table 2-1. Tonal Contrasts in Bai.**

| | Tense | Lax | Nasal Tense | Nasal Lax |
|---|---|---|---|---|
| High Level (55) | Constricted (Harsh Voice) | Unconstricted (Modal Voice) | Constricted (Harsh Voice) | Unconstricted (Modal Voice) |
| Mid (33) | Constricted (Harsh Voice) | Unconstricted (Modal Voice) | Constricted (Harsh Voice) | Unconstricted (Modal Voice) |
| Mid Falling (31) | Constricted (Harsh Voice) | Unconstricted (Breathy Voice) | Constricted (Harsh Voice) | Unconstricted (Breathy Voice) |
| Low Falling (21) | Constricted (Harsh Voice) | N/A | Constricted (Harsh Voice) | N/A |
| Mid Rising (35) | Constricted to Unconstricted (Harsh Voice to Modal Voice) | N/A | N/A | N/A |

The acquisition of Bai by infants necessarily involves refined laryngeal control at the syllabic level. Esling's (2005) model of the vocal tract, depicted in Figure 2.2 below, serves to clarify the phonetic parameters and associated anatomical structures involve in developing laryngeal control. Esling's (2005) model characterizes the larynx as a separate articulator, distinct from the oral vocal tract, that can be controlled in the horizontal and vertical planes, accounting for laryngeal features that can be manipulated at and above the glottis, respectively.

**Figure 2-2. Esling's (2005) model of the vocal tract.**




Building on this model, Edmondson and Esling (2006) have identified six valves that can be controlled within the laryngeal articulator. Valve 1 is responsible for the control of glottal vocal fold adduction and abduction. The relatively passive Valve 2 is responsible for the incursion of the ventricular folds over the vocal folds, a maneuver seen in moderately articulated glottal stops. The activation of Valve 3 produces the sphincteric compression of the arytenoid cartilages and the attached aryepiglottic folds. The activation of this valve is seen, in ascending degrees, in constricted voice qualities such as whisper, creaky voice, and harsh voice, which represent slight, moderate, and strong degrees of activation of this valve, respectively. When Valve 3 is not activated, unconstricted voice qualities such as modal voice, breathy voice, and falsetto can be produced with the action of Valve 1. Valve 4 involves tongue retraction and backwards movement of the epiglottis, a mechanism that is necessary for the production of pharyngeal consonants. The activation of Valve 5 produces raising and lowering of the larynx, which changes vocal tract resonance. Finally, Valve 6 involves inward constriction of the pharynx, such as is seen in the most highly constricted articulations.

To date, Esling's research framework has been applied to the study of whisper in Chinese tones (Gao, 2002; Gao & Esling, 2003) and to voice quality in Japanese anime (Teshigawara, 2003). More recently, this model has been used in studies of vocal tract modeling (Moisik & Esling, 2007; Moisik, 2008) and in studies of infant vocalizations (Bettany, 2004; Esling et al., 2004, 2006, Benner et al., 2007, Grenon et al., 2007). The study reported in this dissertation employs Esling's (2005) model of the vocal tract to provide a clear means of classifying the full range of infant vocalizations on the basis of laryngeal voice quality and to explore adult perceptions of such utterances.

At this time, acoustic research into laryngeal voice quality is not sufficiently developed to support a full instrumental analysis of laryngeal constriction in infant vocalizations. However, previous research has shown that auditory coding, in conjunction with visual inspection of spectrograms, can support reliable classification of infant vocalizations (Buder et al., 2008; Nathani & Oller, 2001; Rvachew et al., 2002). Nor is it yet possible to systematically manipulate infant speech stimuli to produce natural-sounding utterances that vary only in terms of the laryngeal voice quality features discussed in this dissertation. Thus, the ability to conduct controlled perceptual tests that focus exclusively on the role of laryngeal voice quality in adults' perceptions of infant vocalizations is limited. However, given the apparent salience of laryngeal voice quality variations in adults' perceptions of infant crying, it seems likely that adults would be responsive to such features in other infant vocalizations. The results of this study may assist in developing future perceptual studies of laryngeal voice quality that use different methods of controlling for confounding sources of variation in adults' perceptions of laryngeal voice quality.

## 2.4 Summary

This literature review has provided a broad discussion of a range of physiological, psychological, and social factors that may affect the development of infants' vocalizations in the first year of life, as well as adults' perceptions of such vocalizations. First, infants are born with a physiological disposition that favours the production of vocalizations exhibiting laryngeal constriction. Infants are also born with a psychological disposition that is geared towards vocal attunement to their caregivers. This factor, in combination with infants' predisposition to explore their evolving capacities—vocal and otherwise—stimulates infants to explore their phonetic abilities in vocally playful activities that eventually result in the production of utterances that begin to resemble those they hear in their ambient language, mirroring, to some extent, the early attunement to the target language evidenced in studies of infant speech perception. Thus, towards the end of the first year, an increasing proportion of infants' production is syllabic, as reflected in the increase in the production of canonical babbling. In addition, babbling may begin to reflect prosodic features that are salient in the language and/or that are favoured in infants' cultural and linguistic environment.

To date, it is unknown whether infants' use of laryngeal voice quality is systematically related to the use of voice quality features in the infants' ambient language, or whether voice quality features are distributed differently between syllabic versus non-syllabic non-distress vocalizations. This dissertation addresses this gap in the literature by analyzing the development of laryngeal voice quality features among infants who are learning two very different languages: English, which uses laryngeal voice quality primarily for paralinguistic expression, and Bai, which employs laryngeal voice quality contrastively at the syllabic level, as part of its register tone system. Moreover, this study is conducted within a phonetic framework that provides a consistent theoretical perspective on laryngeal voice quality features, based on extensive laryngoscopic observations of the adult larynx.

Finally, given the documented role of adults in selectively reinforcing vocalizations that they perceive to be desirable in their language and/or culture, and discouraging less favoured vocalizations, this dissertation features an exploratory study of adults' perception of laryngeal voice quality in infant vocalizations, with particular focus on the influence of laryngeal constriction in syllabic versus non-syllabic utterances. To explore this topic, the perceptions of adults from two language groups that contrast strongly in their use of laryngeal constriction in speech (English and Bai) will be examined. To my knowledge, no phonetic or linguistic research has focused on adult perceptions of laryngeal voice quality in syllabic and non-syllabic infant vocalizations. This study attempts to address this previously unexplored issue, with a view to better understanding the relationship between infants' production of laryngeal voice quality features and adults' perception of a range of utterances that vary in terms of this feature.

# Chapter 3

# METHODOLOGY

## 3.1 Study 1: Infant Production

### *3.1.1 Participants*

Infant participants were selected from a larger database of recordings made in English-speaking homes in Canada and Bai-speaking homes in China. The larger database included infants who were recorded either longitudinally or cross-sectionally over the first year. Subjects for the present study were selected on the basis of having been recorded at least once during each of the four age periods analyzed in the production analysis: (1) months 1-3; (2) months 4-6; (3) months 7-9; and (4) months 10-12. Of the available infants, those infants who were recorded most frequently during each quarterly age period were selected, so that each month of the first year would be represented in the overall age categories.

Based on these criteria, four Bai infants (2 female, 2 male) and four English infants (3 female, 1 male) were selected for the present study. While not all infants were recorded each month, the data presented in this dissertation include at least two infants per language per month, and each quarter includes the production of four infants per language. The number of infant recordings, per quarter for each language and for each infant in the study, is indicated in parentheses in Table 3-1 below.

**Table 3-1. Recording sessions per infant, by age group.**

| Age Group | Bai Infants | English Infants |
|---|---|---|
| Months 1-3 | Infant B-1 (8)<br>Infant B-2 (3)<br>Infant B-3 (3)<br>Infant B-4 (2) | Infant E-1 (2)<br>Infant E-2 (2)<br>Infant E-3 (1)<br>Infant E-4 (5) |
| Months 4-6 | Infant B-1 (4)<br>Infant B-2 (3)<br>Infant B-3 (3)<br>Infant B-4 (1) | Infant E-1 (10)<br>Infant E-2 (2)<br>Infant E-3 (3)<br>Infant E-4 (5) |
| Months 7-9 | Infant B-1 (5)<br>Infant B-2 (3)<br>Infant B-3 (3)<br>Infant B-4 (2) | Infant E-1 (12)<br>Infant E-2 (1)<br>Infant E-3 (3)<br>Infant E-4 (1) |
| Months 10-12 | Infant B-1 (6)<br>Infant B-2 (3)<br>Infant B-3 (2)<br>Infant B-4 (2) | Infant E-1 (12)<br>Infant E-2 (2)<br>Infant E-3 (1)<br>Infant E-4 (1) |

### *3.1.2 Recording procedure*

The four Bai infants in the study were videorecorded using a Sony DCR-HC26 digital video camera recorder with an integrated microphone. Infants were recorded at 16-bit and at a sampling rate of 44,000 samples per second. Infants were filmed in spontaneous interaction with a caregiver. While the caregiver was usually the infant's mother, in some cases, the filmed caregiver was the father or grandmother; with some infants, the filmed caregiver varied from session to session. The camera was directed at the infant and held at a reasonable distance so as to not distract the infant from the caregiver while ensuring proper capture of the sounds produced by the infant. Recordings were made in the infants' homes and lasted approximately 30 minutes each.

Two of the four English infants included in the study were recorded in the manner described above for the Bai infants. The third English infant was audiorecorded using an

Audio-Technica® PRO 7a unidirectional condenser microphone and a TASCAM DA-P1 Digital Audio Tape (DAT) recorder at 16-bit, 22,050 samples per second. The fourth English infant was audiorecorded on a a Uher 4000 portable tape recorder at 7.5 inches-per-second with a a Sony electret condenser microphone held at about 10 to 18 inches from the infant's mouth. The analog data from the latter recordings were captured and digitized at 16-bit, 44,000 samples per second, using SoundForge 8.0.

This procedure resulted in 58 hours of recorded data for the infants included in this study. The recordings were transferred to a personal computer for segmentation and analysis.

### *3.1.3 Selection and Segmentation of Utterances*

The audio and video recordings were segmented into vocalizations and stored as separate files using SoundForge 8.0. With the exception of crying, all sounds produced by the eight infants in the recording sessions were selected for this analysis, including sounds that have been labeled as cooing, grunting, coughing, squealing, or babbling in previous research. Sounds that could not be judged accurately (e.g. because of background noise) were excluded from analysis. For the purposes of this study, a vocalization was any non-crying sound lasting at least 500 msec and separated from other sounds by at least 2 seconds of silence. To ensure that the audio and video recordings in the study would be judged according to the same auditory criteria described below, the audio portion of the video recordings was split from the video file and saved separately.

This procedure resulted in a database of 4,153 utterances (2024 Bai; 2129 English), broken down by age group and language as represented in Table 3-2.

**Table 3-2. Number of utterances by language and age group.**

| Age Group | Bai Infants | English Infants |
|---|---|---|
| Months 1-3 | 445 | 316 |
| Months 4-6 | 351 | 762 |
| Months 7-9 | 619 | 495 |
| Months 10-12 | 609 | 556 |

To ensure an equal number of samples from within each age category for each language for purposes of statistical analysis, it was necessary to limit the number of utterances analyzed from the total utterances available. Because there were relatively few utterances for the English infants during Months 1-3 (316), and for the Bai infants in Months 4-6 (351), 300 utterances for each age group and language were selected for further analysis, for a final database of 2400 utterances, 1200 for each language group. In selecting the utterances to be analyzed, privilege was given to ensuring that each month within the age category was represented equally. Thus, 100 sounds from each of the 12 months studied were selected for the final database. Within each month, sounds were selected to ensure that an equal proportion of sounds produced by each of the infants who were recorded during that month were included. For example, for Month 1, the Bai sample included recordings from four infants. Thus, 25 sounds from each infant were selected for that month, except when an infant did not produce 25 utterances. In the latter cases, the balance of the 25 sounds originally allotted to that infant was distributed among the remaining infants. If an infant was recorded more than once during a given month, then an equal proportion of the sounds selected for that infant, for that month, were chosen from the available recordings. For instance, recordings of two English infants were available for Month 2, so 50 utterances were chosen for each infant for that month. However, one of the babies was recorded twice, so 25 sounds were selected from one recording, and 25 from the other, to reduce the impact of session effects.

Overall, the selection of utterances was made according to the following criteria, in order of priority: (1) to ensure an equal proportion of sounds for each language, for each quarterly age category; (2) to ensure an equal proportion of sounds for each month within a larger quarterly period (i.e., 100 sounds from each of months 10, 11, and 12 for the period Months 10-12); (3) to ensure an equal proportion of sounds from each infant within a given month, subject to the available number of recordings; and (4) to ensure an equal proportion of sounds from each available session within a given month, in cases where an infant was recorded more than once in a given month.

### *3.1.4 Auditory Coding*

The 2400 infant vocalizations selected for categorization were classified using auditory analysis, supplemented by examination of wide-band spectrograms. Utterances were classified according to laryngeal quality and utterance type.

#### *3.1.4.1 Voice Quality Categories*

For laryngeal quality, utterances were divided into three broad categories: constricted, unconstricted, and dynamic. Constricted utterances were those produced with harsh voice, creaky voice, whispery voice, or whisper. The spectrograms below are examples of the types of utterances that were classified in the present study as constricted. Figures 3-1 below is a spectrogram of a constricted utterance produced with harsh voice. This vocalization is extremely harsh in the middle portion of the utterance, as evidenced by the distribution of high energy throughout the spectrum, including the portion above 3000 Hertz, which is indicated by the horizontal line running through the spectrogram .

**Figure 3-1. Spectrogram of constricted utterance with harsh voice.**

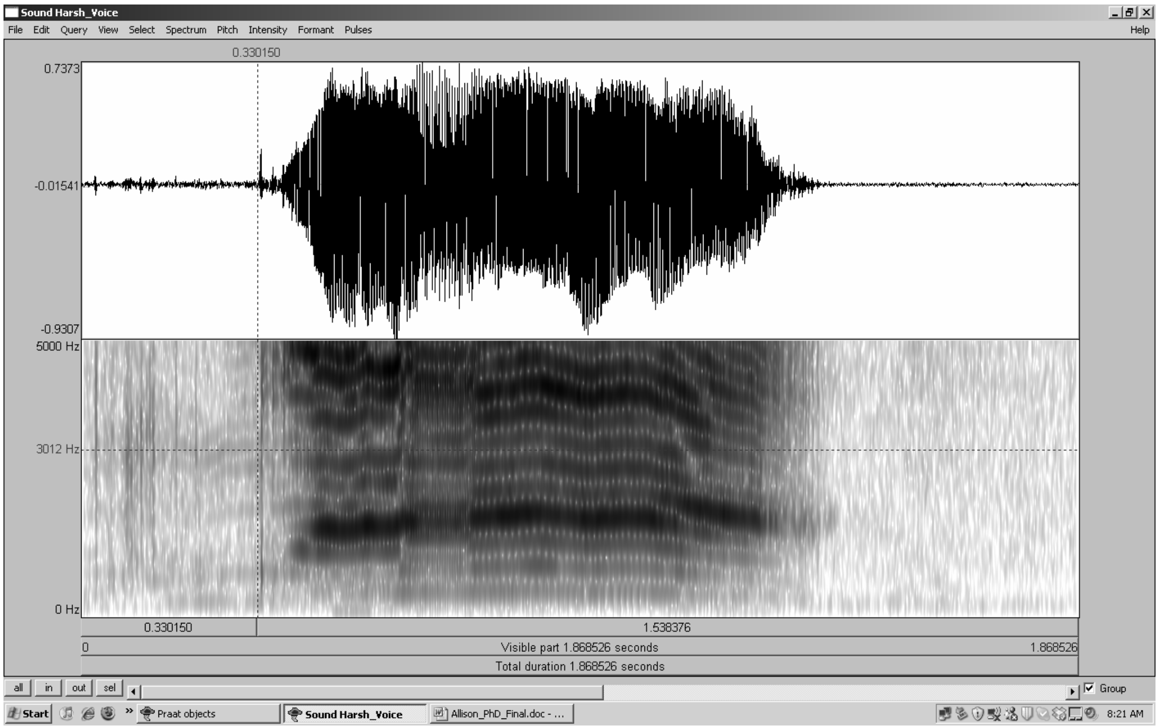



Figure 3-2 below is a spectrogram of a constricted utterance produced with creaky voice. The creaky voice is evident from the widely spaced vertical striations distributed throughout the spectrum.

**Figure 3-2. Spectrogram of constricted utterance with creaky voice.**

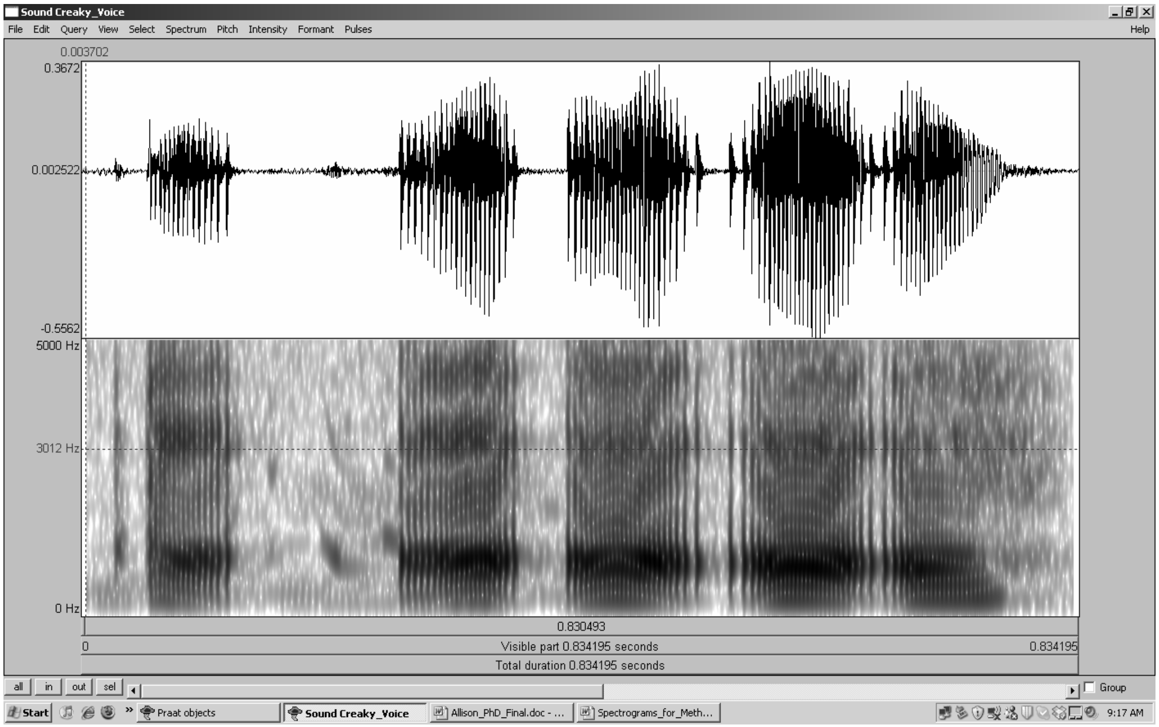
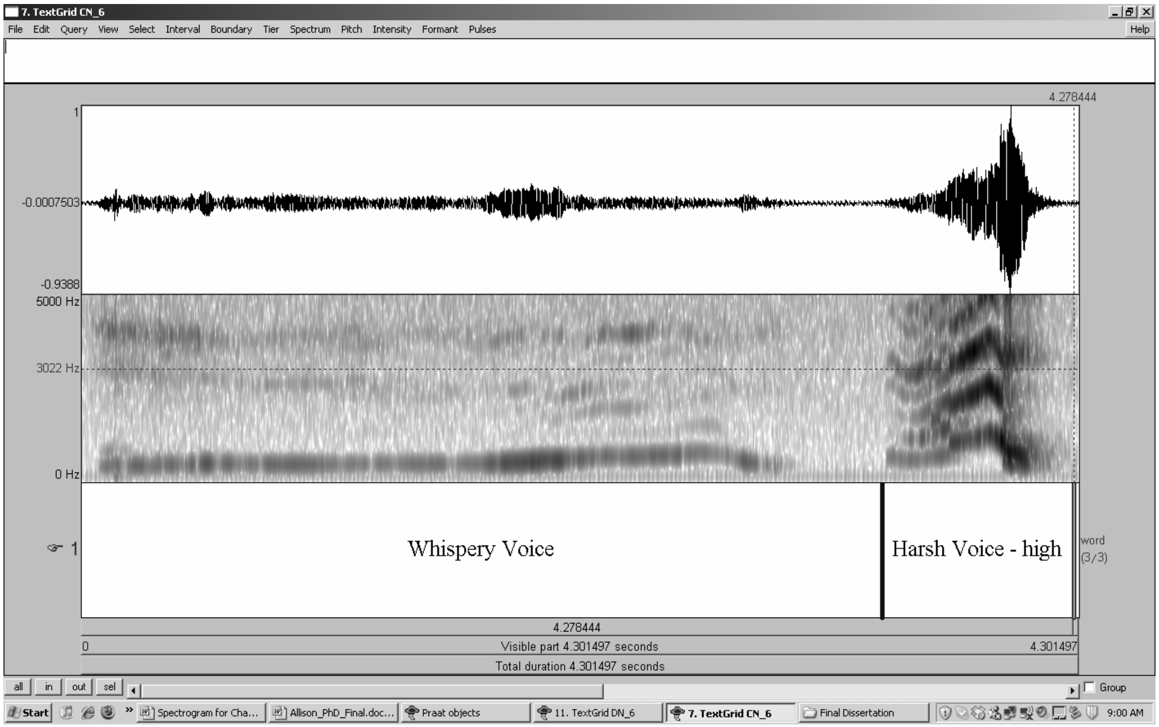


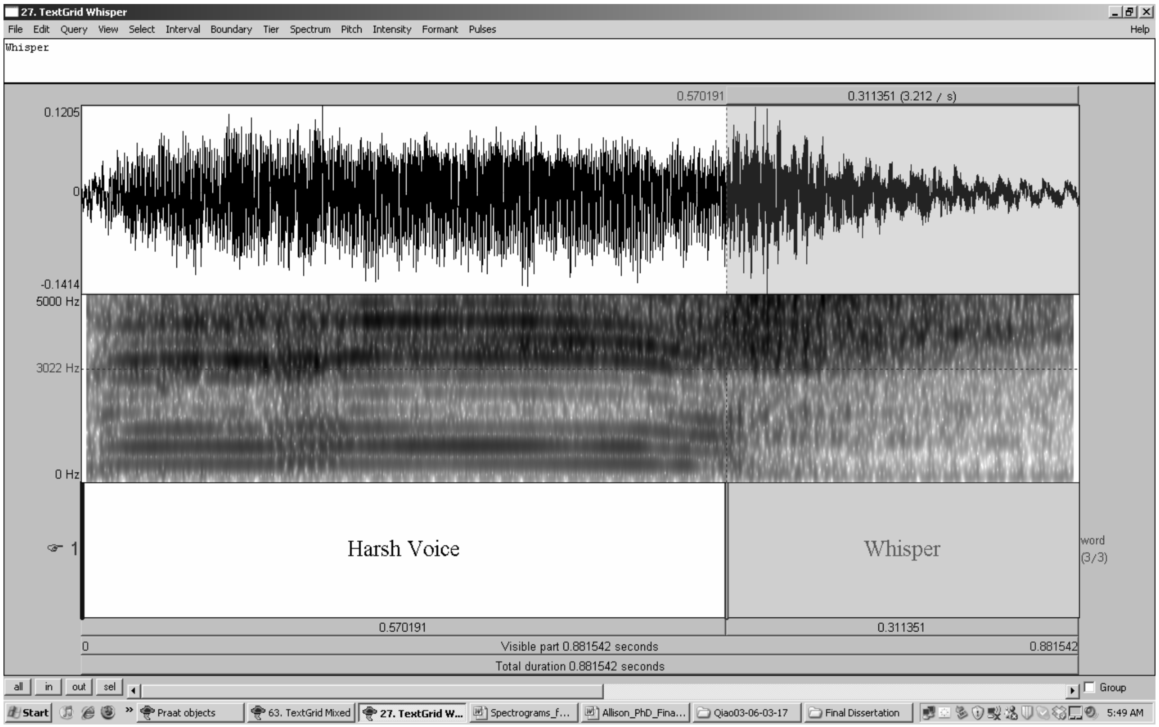

Figure 3-3 below is an example of a constricted utterance; the first part of the sound is produced with whispery voice. This quality is evident in two acoustic features displayed in the spectrogram. First, there is evidence of voicing in the bottom part of the spectrum. Second, there is a band of noise in the upper part of the spectrum, above 3000 Hertz, indicating the presence of laryngeal constriction.

**Figure 3-3. Spectrogram of constricted utterance with whispery voice.**




Finally, Figure 3-4 is an example of a constricted utterance produced with whisper. The whispered portion of the utterance is towards the end of the vocalization, as labelled on the spectrogram. The whisper quality is evident from the absence of a voicing bar at the bottom of the spectrogram, and from the presence of a band of noise in the upper portion of the spectrum, above 3000 Hertz.

**Figure 3-4. Spectrogram of constricted utterance with whisper.**




Unconstricted utterances were those produced with modal voice, breathy voice, or falsetto. The spectrogram shown in Figures 3-5 below is an example of an unconstricted utterance produced with modal voice. In the spectrogram, the unconstricted quality of this utterance is evident in the relative absence of energy in the mid- to upper portion of the spectrum, above 3000 Hertz, compared to the presence of energy in this region in the spectrograms shown for constricted utterances in Figures 3-1, 3-2, 3-3, and 3-4 above (the energy distributed throughout the spectrum in the spectrogram below is due to background noise, not the infant’s production).

**Figure 3-5. Spectrogram of unconstricted utterance with modal voice.**

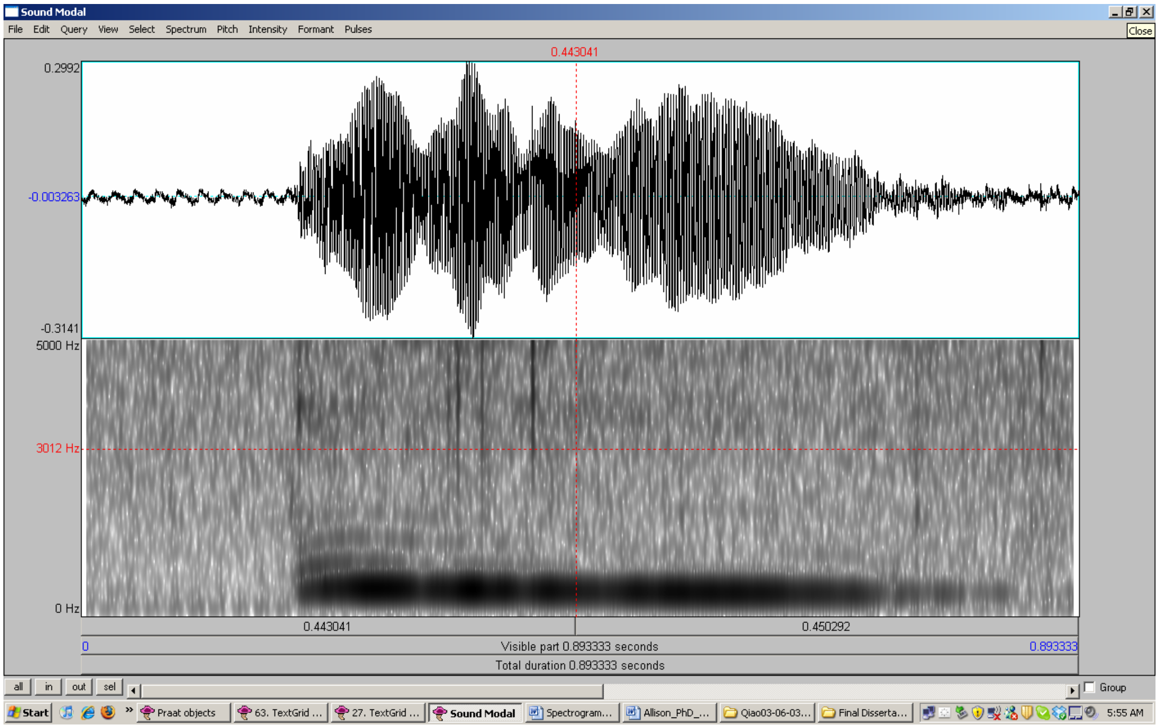
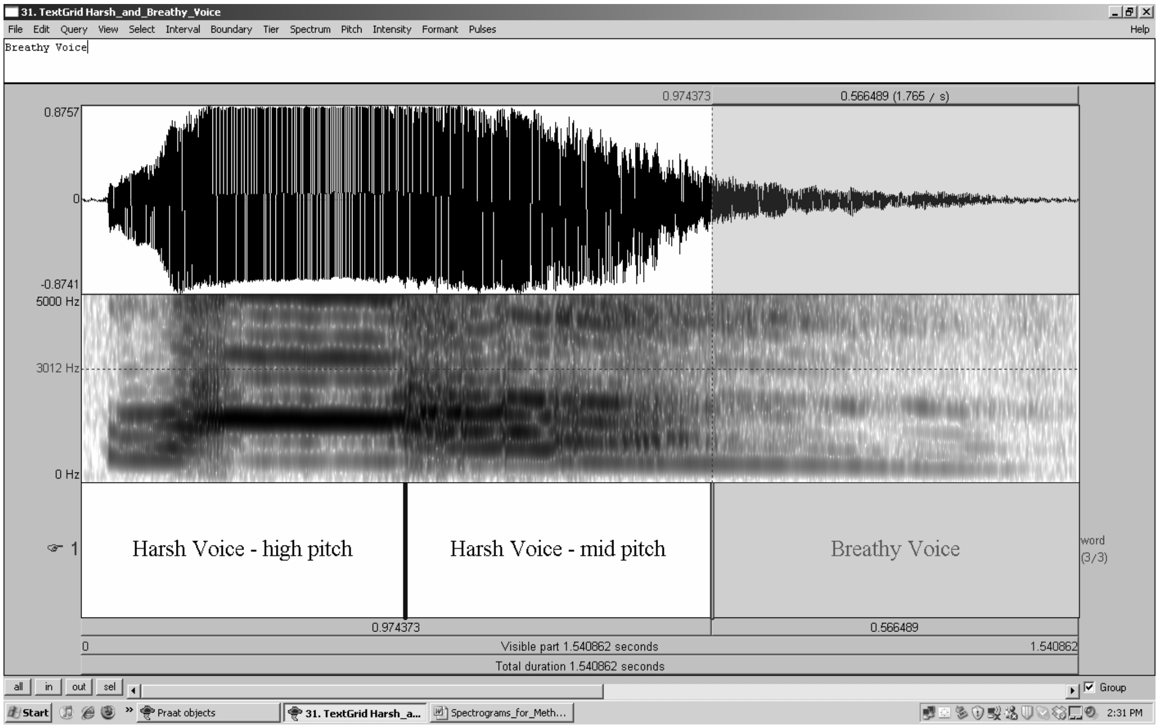


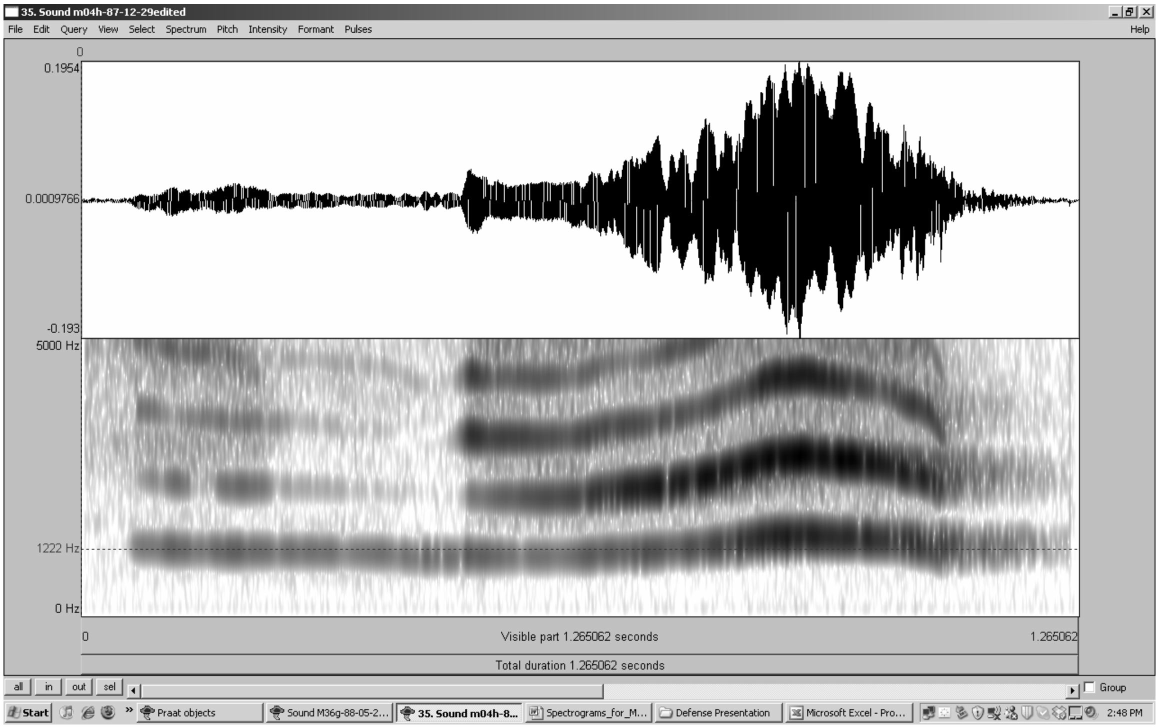

Figure 3-6 below is an example of breathy voice. In the corpus of data used for the present study, breathy voice was seldom produced in isolation. Thus, the breathy quality illustrated below occurs at the end of an utterance that starts out in high-pitched harsh voice and moves into mid-pitched harsh voice, before becoming breathy as pitch declines. As indicated in the annotation of the spectrogram, the breathy portion  is at the end of the utterance, and is evident in the voicing bar at the bottom of the spectrogram and the distribution of low-amplitude noise throughout the spectrum.

**Figure 3-6. Spectrogram of dynamic utterance with harsh and breathy voice.**



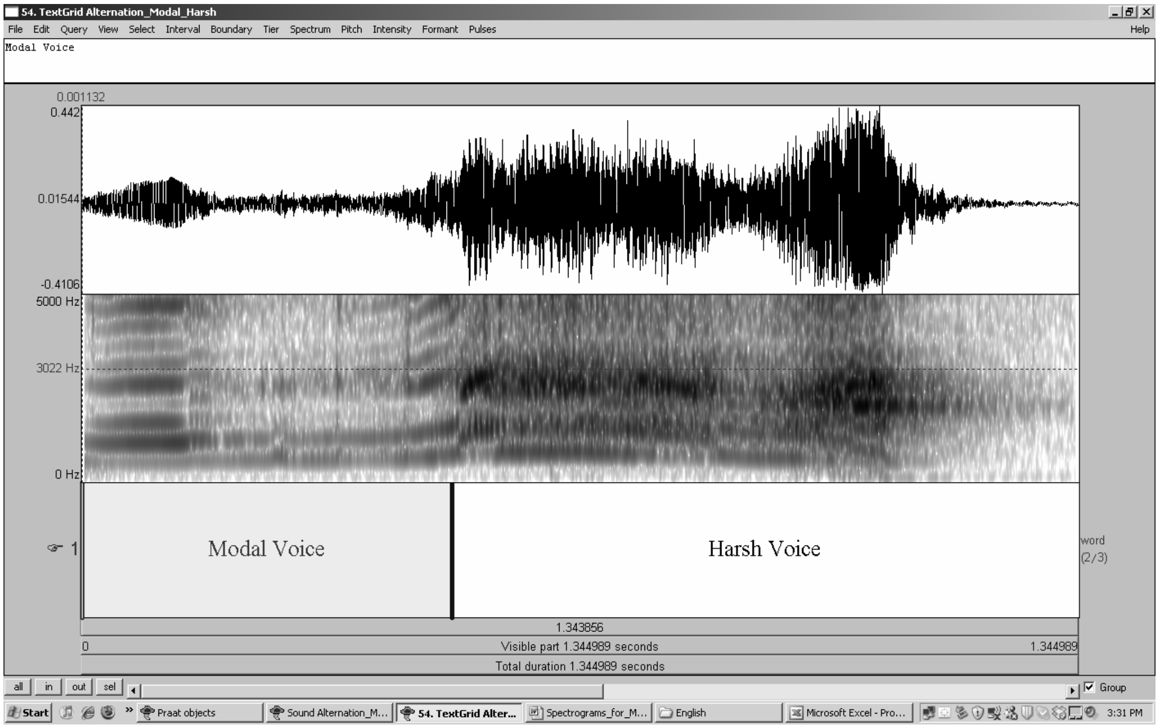

The spectrogram shown in Figure 3-7 below is an example of an unconstricted utterance produced with falsetto. The falsetto quality is evident in the widely spaced harmonics throughout the spectrum (in high-pitched utterances like this, the fundamental frequency often gives the appearance of being a formant).

**Figure 3-7. Spectrogram of unconstricted utterance with falsetto.**




Dynamic utterances were those that involved at least one alternation between a constricted and an unconstricted voice quality setting within a single vocalization. Figure 3-6 above, which illustrates an alternation between harsh voice and breathy voice, is an example of a dynamic utterance. Figure 3-8 below is another example of a dynamic utterance. In this utterance, the utterance starts out in modal voice and ends in harsh voice, as seen in the increase in energy in the upper portion of the spectrum towards the end of the utterance.

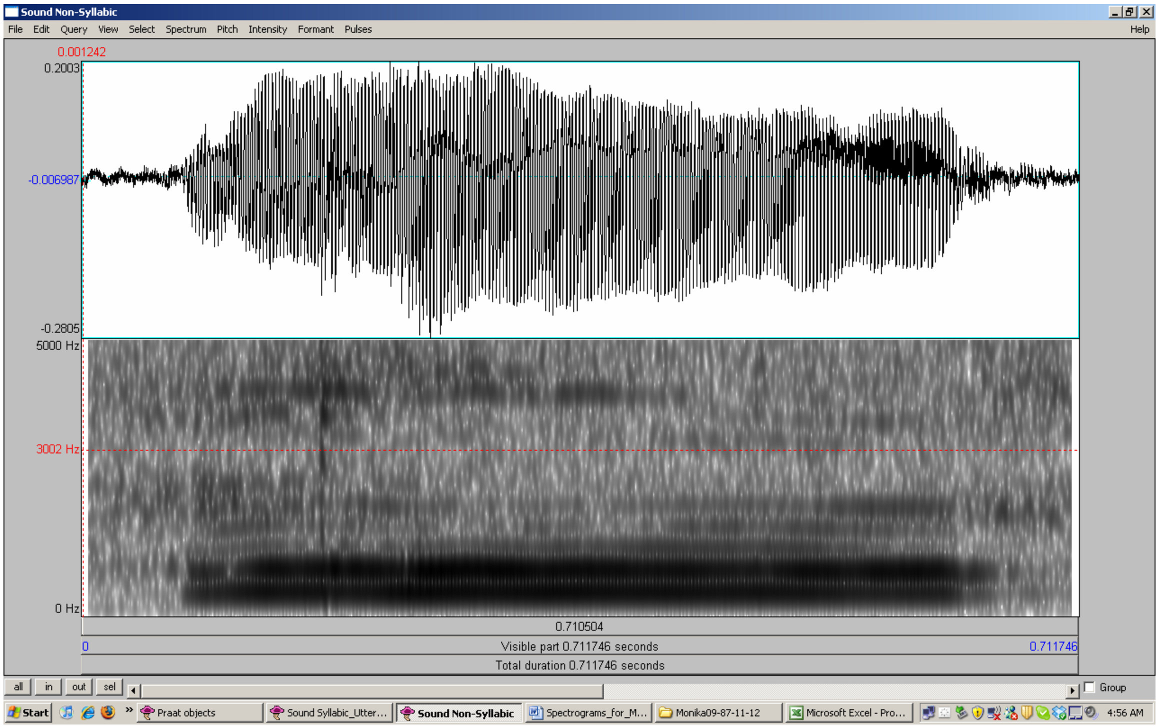
**Figure 3-8. Spectrogram of a dynamic utterance.**




*3.1.4.2 Utterance Categories*

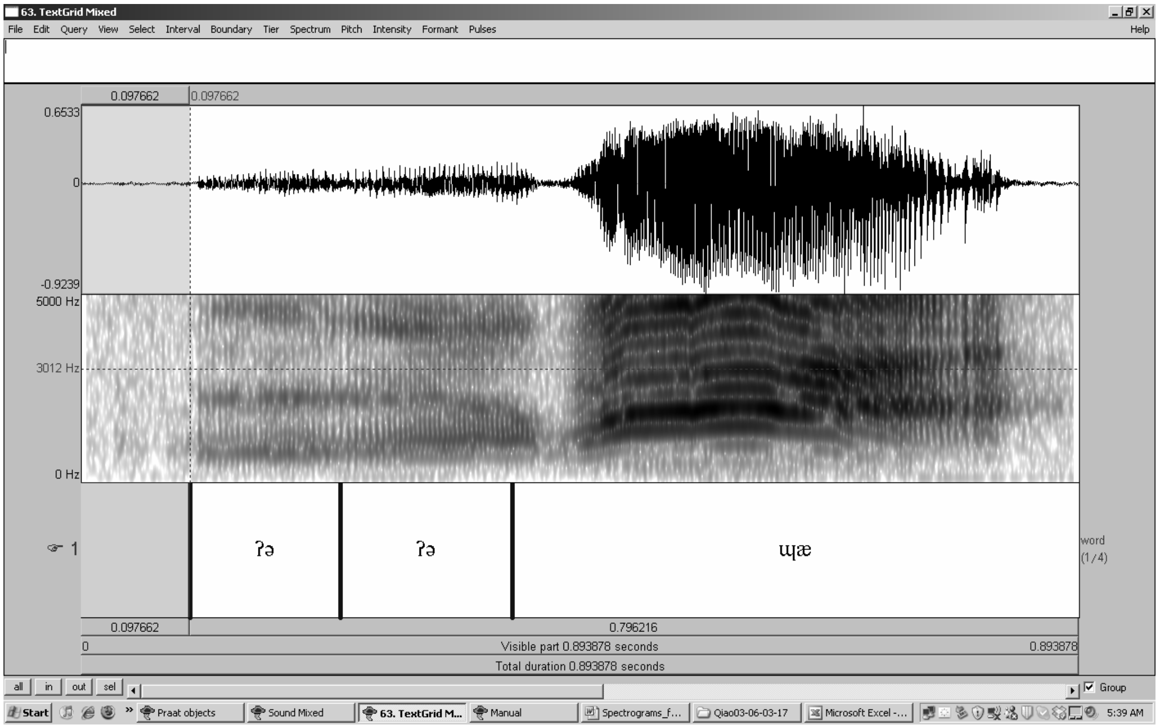
In terms of utterance type, all vocalizations were divided into three broad categories: non-syllabic, mixed, and syllabic. In the present study, non-syllabic utterances were vocalic utterances that did not include oral consonants. Figure 3-9 is an example of a non-syllabic utterance. This utterance is a prototypical non-syllabic utterance in that it features an uninterrupted stretch of voicing. Some non-syllabic utterances consist of a series of short vocalic utterances, sometimes including laryngeal consonants, including glottal and pharyngeal stops and fricatives. However, in contrast to mixed and syllabic utterances, illustrated in Figures 3-9 and 3-10 below, non-syllabic utterances are not broken up into shorter parts by the production of oral consonants.

**Figure 3-9. Spectrogram of a non-syllabic utterance.**



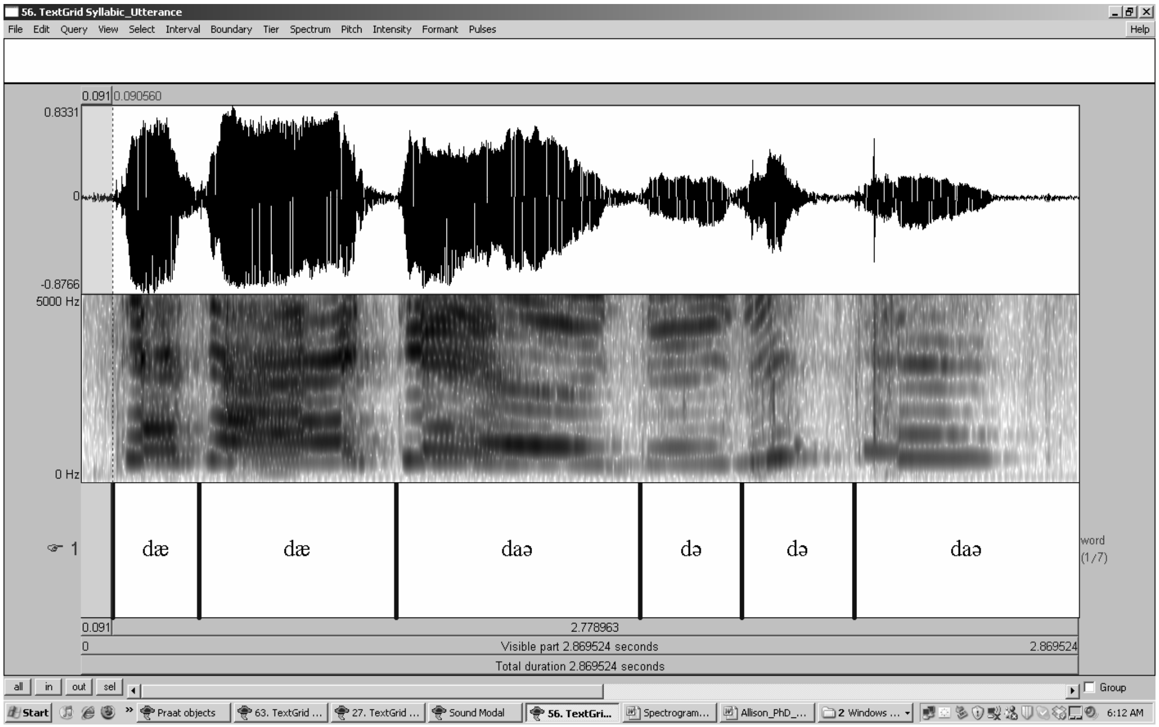

Within this study, the category of mixed utterances was created to distinguish purely vocalic utterances and utterances with laryngeal consonants from those that included some constriction within the oral vocal tract, such as a glide, approximant, or trill. Figure 3-10 is a spectrogram of a mixed utterance. As shown in the annotation of the spectrogram, the first part of the utterance features a sequence of glottal syllables. Towards the middle of the utterance, where there is a dampening of energy in the spectrum, the infant produces a velar approximant. However, there is no cessation of voicing and the production of the vowel-like oral consonant does not break the utterance into two distinct syllables.

**Figure 3-10. Spectrogram of a mixed utterance.**




Finally, syllabic utterances included all CV(C) utterances typically considered to be babbling, whether monosyllabic, reduplicative, or variegated. CV(C) utterances produced with laryngeal consonants (pharyngeal or glottal) were not grouped into this category, unless they were produced as part of a babbling sequence that included at least one oral consonant. Figure 3-11 illustrates a voiced alveolar babbling sequence that is classified as syllabic within this study. The vocalic portion of the utterance is broken up by a series of silences and releases where the voiced alveolar stops are produced, as shown in the annotation of the spectrogram.

**Figure 3-11. Spectrogram of a syllabic utterance.**




*3.1.4.3 Reliability*

Once the utterances were classified by the author using the categories described above, the classifications were checked against the auditory judgments made by a group of trained phoneticians on the same utterances in a separate database compiled as part of the Infant Speech Acquisition (InSpA) Project at the Department of Linguistics, University of Victoria. The judgments made for this study were found to correspond to those made on the same utterances in the InSpA Project 87% of the time, with most areas of disagreement relating to the classification of: (a) breathy voice versus whispery voice; and (b) modal voice versus mild to moderate harsh voice, such as would be produced with larynx raising and no trilling of the aryepiglottic folds or aperiodic voicing. In these cases, the author re-examined the utterances with a trained phonetician until agreement was reached, or selected a different utterance from the available infant utterances used in this study. No systematic acoustic measurements were made of the infant utterances because the recordings differed too widely in quality to make such measurements reliable.

### *3.1.5 Statistical Analysis*

The results of the auditory classification of infant utterances were analyzed in SPSS 17.0 using chi-square and Cramer's V to test the strength of the association. The purpose of the statistical analysis was to determine whether the distribution of laryngeal voice quality features differed significantly according to age, utterance type, and language background. Results of the statistical analysis for utterance type by age and language are reported in Section 4.2.4. Results of the statistical analysis for laryngeal voice quality by age, utterance type, and language are presented in Section 4.3.4.

## 3.2 Study 2: Adult Perception

### *3.2.1 Participants*

Forty adult native speakers of Bai (24 female, 16 male), ranging in age from 20 to 75, and 40 adult native speakers of English (19 female, 21 male), ranging in age from 22 to 77 participated in the listening task. Bai participants were recruited by the Project Coordinator responsible for recording Bai infants in Jianchuan City, Yunnan, China for the Infant Speech Acquisition (InSpA) Project, based at the University of Victoria. All Bai participants were native speakers of the Jianchuan dialect of Bai, which features laryngeal register tone contrasts. English participants were recruited by the author by placing an advertisement in a community centre in Victoria, Canada and by distributing a flyer to residents in selected neighbourhoods of Victoria. All English participants were native speakers of Canadian English. None of the Bai or English participants reported any speech or hearing impairments, and none of the participants had received formal training in linguistics or phonetics.

### *3.2.2 Task Procedure*

Participants rated 36 randomly ordered infant vocalizations according to how important they considered the utterances to be in learning to speak their native languages (English or Bai). Participants rated the 36 infant vocalizations on a five-point scale, where ‘1’ was used to designate sounds that were unimportant in learning to speak their native language, and ‘5’ was used to mark sounds that listeners considered very important in learning to speak the native language. Prior to taking the test, participants were reminded that all of the sounds in the listening task were produced by prelinguistic infants, and to bear this fact in mind in rating the utterances they were about to hear. Participants were also asked to try to use all five points on the rating scale.

Participants completed the listening task on a laptop computer and listened to the infant vocalizations via headphones. The experiment was designed and conducted using E-prime 2.0. The computerized template for the experiment started with written instructions in Chinese for the Bai participants, and in English for the English participants, and a brief training session in which participants rated five infant vocalizations representing the types of utterances that would be included in the perception test. After the playing of each sound, the screen displayed a five-point scale, labeled to illustrate the meaning associated with each end of the scale (i.e., “1” was labeled “not important at all” and “5” was labeled “very important”). Listeners registered their evaluation of each sound by pressing the appropriate number key. After entering their rating, a screen appeared instructing them to proceed to the next sound by pressing the space bar. Most listeners took approximately 15 minutes to complete the listening task.

Following the listening task, the participants were interviewed for approximately five minutes. Participants were asked to provide information on their age, languages spoken, and their experience with infants. Participants were also asked to discuss the factors that they believed contributed to their own low and high ratings, respectively. Interviews with the Bai participants were conducted by an interviewer in Mandarin and were filmed with a digital camcorder by the author. The interviewer, a trained phonetician and native speaker of Mandarin, translated the participants’ comments into English. Interviews with the English participants were conducted in English by the author. Participants’ responses were recorded by the author on a written form.

After completing the interview, participants received a small gift. Bai participants received a small honorarium and a book of postcards from Victoria, Canada. English participants received a voucher for a beverage at a local coffee shop.

The listening task was piloted with five native English speakers, who reported no difficulty in completing the task or in using all points on the scale. While it was not possible to pilot the test with phonetically naïve speakers of Bai, two trained phoneticians who were native speakers of Bai completed the listening task two days prior to the administration of the test in Jianchuan as a test of the viability of the listening task in a non-English context. These speakers reported no difficulty in completing the task, and their data were discarded.

### *3.2.3 Infant Speech Stimuli*

The infant speech samples used in the listening task were excerpted from the infant speech samples included in the production component of the present study. Infant speech samples were selected and edited to contain a minimum of background noise and to exclude the voices of adults or other children before being normalized for volume. Because the source recordings often included loud background noise and/or the voices of caregivers and other children, these criteria restricted the available choices and the ability to control for the phonetic features that were the specific focus of investigation—factors which must be borne in mind when interpreting the results of this study.

Eighteen of the vocalizations were produced by Bai infants, and 18 were produced by English infants. These speech samples were reviewed by a native speaker of English to ensure that they did not feature any English words. The samples were also evaluated by a native speaker of Bai, who is also a trained phonetician. This speaker reported that while the samples included every known tonal paradigm in Bai, they did not inadvertently include any lexical items in Bai.

### *3.2.4 Phonetic Categories*

The infant vocalizations were selected to be representative of the five broad phonetic categories employed in the production component of the present study: syllabic, non-syllabic, constricted, unconstricted, and dynamic. These broad categories then yielded six smaller sub-categories: constricted syllabic and constricted non-syllabic, unconstricted syllabic and unconstricted non-syllabic, and dynamic syllabic and dynamic non-syllabic. In each category and sub-category, an equal number of Bai and English infant vocalizations were included. A more detailed description of the samples selected from within each category and sub-category follows below.

The syllabic vocalizations included in the listening task were all examples of babbling produced by Bai and English infants from the two later age categories examined in the production component of this study: months-7-9 and months 10-12. The consonants and vowels in these vocalizations ranged widely, but most vowels were central monophthongs and all the consonants were bilabial, alveolar, and/or velar. Most of the babbling was canonical (i.e., featuring a repetition of the same CV pattern), but some babbling sequences varied in place and/or manner of articulation. Eighteen of the 36 vocalizations in the listening task were syllabic. The length of the 18 syllabic utterances included in the listening task ranged from 2.64 seconds to 5.36 seconds, with a mean length of 3.55 seconds.

The 18 non-syllabic utterances included in the listening task were all vocalic, though some included a glottal stop or fricative. The dynamic non-syllabic utterances in the task were representative of the non-syllabic sub-category “mixed,” in which utterances contain a small constriction in the oral vocal tract, typical of a glide or approximant. All of the non-syllabic utterances included in the listening task were produced by infants in the same age ranges as for the syllabic utterances (7-9 months and 10-12 months), although many of these types of sounds are produced by much younger infants. The 18 non-syllabic utterances ranged in length from 1.64 to 6.26 seconds, with a mean length of 4.16 seconds.

The constricted utterances in the listening task included harsh voice, creaky voice, and/or occasional traces of whisper or whispery voice. The utterances produced with harsh voice varied considerably in pitch, ranging from low-pitched harsh voice to very high-pitched harsh voice. The vocalizations also featured degrees of harshness, ranging from very harsh voice with audible aryepiglottic trilling, to the slightly constricted voicing associated with larynx raising. Some of the samples were representative of a single constricted voice quality (e.g., harsh voice), while others featured more than one constricted voice quality (e.g., harsh voice and creaky voice). Twelve of the 36 vocalizations were constricted. Of these 12, 6 were syllabic and 6 were non-syllabic.

The unconstricted vocalizations in the listening task featured modal voice, breathy voice, and/or falsetto. As with the constricted vocalizations, the unconstricted vocalizations varied in pitch and quality, but all the unconstricted vocalizations included modal voicing. Some were modal throughout, while others included modal voice in combination with breathy voice and/or falsetto. Twelve of the 36 vocalizations were unconstricted. Of these, six of the utterances were syllabic and six were non-syllabic.

The dynamic vocalizations included in the perception test alternated between constricted and unconstricted voice quality settings. Twelve of the 36 vocalizations were dynamic; half of these were syllabic, and the other half were non-syllabic. As with the categories discussed above, the dynamic vocalizations exhibited considerable variability. The non-syllabic variants were typical of the types of vocalizations infants produce during vocal play, which often involve sudden changes in laryngeal voice quality, often in concert with wide swings in pitch. Some of the syllabic variants feature two closely connected babbling sequences that are distinguished by a sudden change in laryngeal voice quality (e.g., *bababa* produced in high harsh voice, followed immediately by *bababa* produced in modal voice). Other samples feature a more unified babbling sequence in which one or more syllables is produced with a different laryngeal voice quality (e.g., *babababababa* uttered mostly in modal voice, but with the second and fifth syllables uttered in creaky or harsh voice).

**Table 3-3. Number of utterances by phonetic category in perception test.**

| | Syllabic | Non-syllabic | Total |
|---|---|---|---|
| Constricted | 6 | 6 | 12 |
| Unconstricted | 6 | 6 | 12 |
| Dynamic | 6 | 6 | 12 |
| Total | 18 | 18 | 36 |

### *3.2.5 Statistical Analysis*

The results of the listening task were analyzed using a mixed ANOVA design in SPSS 17.0 to identify the interaction of laryngeal voice quality and/or utterance type in affecting the ratings of infant vocalizations, within and across the two language groups and by gender. Mixed ANOVA results for laryngeal voice quality are reported in

Sections 5.1.1, 5.1.2, and 5.1.3; results for utterance type are reported in Sections 5.2.1, 5.2.2, and 5.2.3. A one-way ANOVA was performed on each of the sub-categories constricted non-syllabic, dynamic non-syllabic, unconstricted non-syllabic, constricted syllabic, dynamic syllabic, and unconstricted syllabic to test the significance of differences in sub-category ratings within and across language groups. Statistical results for non-syllabic sub-categories are presented in Section 5.3.1, and results for syllabic sub-categories are reported in Section 5.4.

### *3.2.6 Qualitative Analysis*

The results of the participant interviews conducted following the listening task were reviewed and coded to identify factors that affected participants' ratings of the infant speech samples. Given the considerable variability of the utterances included in the listening task, and the exploratory nature of the perception experiment, the purpose of this analysis was to aid in the interpretation of the results of the listening task and to identify factors that could contribute to more controlled perceptual studies of infant vocalizations in the future.

# Chapter 4

# PRODUCTION RESULTS

This chapter presents and discusses the results of the first part of the present study, the analysis of laryngeal voice quality features in the production of English and Bai infants in the first year of life. In Section 4.1, I present general findings on the development of laryngeal voice quality features in the vocalizations of Bai and English infants, independent of utterance type. In Section 4.2, I describe the production of difference utterance categories across the first year by infants of both languages, independent of laryngeal voice quality features. Section 4.3 features an analysis of the distribution of laryngeal voice quality features in different utterance types. In Section 4.4, I conclude this chapter with a summary and discussion of the main findings of the production component of the present study.

## 4.1 Laryngeal Voice Quality Features

In this section, the distribution of laryngeal voice quality features for Bai and English infants over the first year is presented. In Section 4.1.1, I outline the development of constricted voice quality settings. In Section 4.1.2, I describe the incidence of dynamic utterances. In Section 4.1.3, I present the evolution of unconstricted utterances in the first year. In Section 4.1.4, I summarize the findings presented in Sections 4.1.1, 4.1.2, and

4.1.3, highlighting universal and language-specific aspects of the development of laryngeal voice quality features in the first year of life.

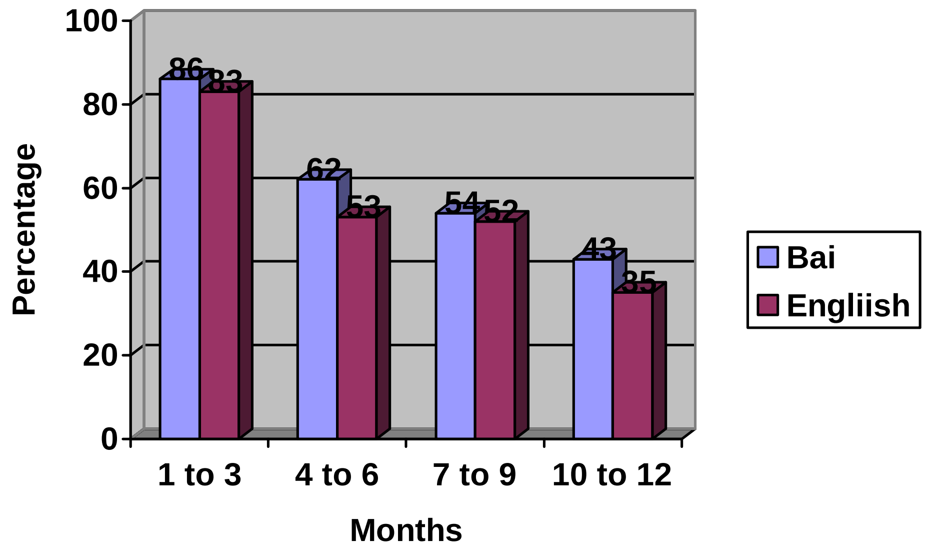
### *4.1.1 Constricted Utterances*

As shown in Table 4-1 and Figure 4-1 below, the overall pattern in the development of constricted utterances is the same for English and Bai throughout the first year of life. While English infants produce slightly fewer constricted utterances in every age period analyzed, and exhibit steeper declines in the production of constricted utterances in months 4-6 and 10-12, the overall pattern in both languages is one of steady linear decline. As illustrated in Figure 4-1, in months 1-3, 86% of Bai infant utterances are constricted. This percentage falls to 62%, 54%, and 43% in months 4-6, 7-9, and 10-12, respectively. In months 1-3, 83% of English infants' utterances are constricted. In months 4-6 and 7-9, the proportion of constricted utterances falls to 53% and 52% respectively, before falling to 35% in months 10-12.

**Table 4-1. Number of constricted utterances of English and Bai infants.**

| Age | Bai Infants (n=734) | English Infants (n=666) |
|---|---|---|
| Months 1-3 | 257 | 248 |
| Months 4-6 | 187 | 158 |
| Months 7-9 | 161 | 155 |
| Months 10-12 | 129 | 105 |

**Figure 4-1. Percentage of constricted utterances of English and Bai infants.**



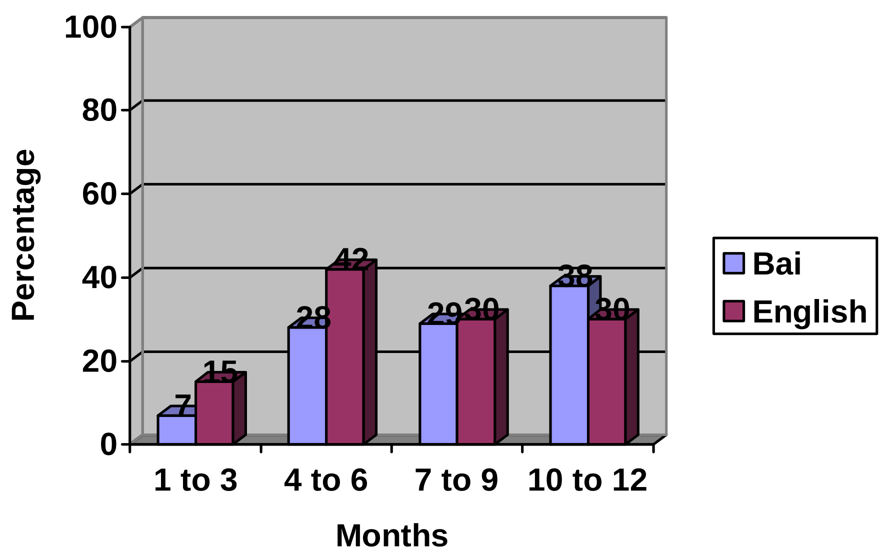

### *4.1.2 Dynamic Utterances*

Table 4-2 and Figure 4-2 illustrate the development of dynamic utterances over the four age periods studied for Bai and English utterances. The incidence and distribution of dynamic utterances appears to differ slightly for English and Bai infants. Overall, English infants produce fewer dynamic utterances than Bai infants (261 for the English infants, versus 308 for the Bai infants). Among English infants, dynamic utterances appear in significant numbers in months 1-3 and increase dramatically in months 4-6. In months 7-9 and 10-12, the incidence of dynamic utterances declines somewhat among the English infants, although these utterances still comprise 30% of the infants' vocalizations. Among the Bai infants, the incidence of dynamic utterances increases significantly in months 4-6, though not to the extent seen among the English infants. Among the Bai infants, the proportion of dynamic utterances holds steady in months 7-9, and then increases in months 10-12. Thus, among the English infants, the incidence of dynamic utterances appears to rise dramatically before slightly declining, while among the Bai infants, the incidence of dynamic utterances rises steadily over the course of the year.

**Table 4-2. Number of dynamic utterances, English and Bai infants.**

| Age | Bai Infants (n=308) | English Infants (n=261) |
|---|---|---|
| Months 1-3 | 22 | 45 |
| Months 4-6 | 85 | 127 |
| Months 7-9 | 88 | 89 |
| Months 10-12 | 113 | 89 |

**Figure 4-2. Percentage of dynamic utterances, English and Bai infants.**




### *4.1.3 Unconstricted Utterances*

Table 4-3 and Figure 4-3 below illustrate changes in the incidence of unconstricted vocalizations over the first year of life for English and Bai infants. For infants from both languages, the pattern is one of linear increase—the opposite of that found for the development of constricted utterances, which exhibit a pattern of linear decline. In months 1-3, only 7% of the Bai infants' utterances are unconstricted. The percentage of unconstricted utterances rises to 10%, 16%, and 19% in months 4-6, 7-9, and 10-12, respectively. Among English infants, the incidence of unconstricted utterances is 2%, 5%, 22%, and 35% in months 1-3, 4-6, 7-9, and 10-12, respectively.

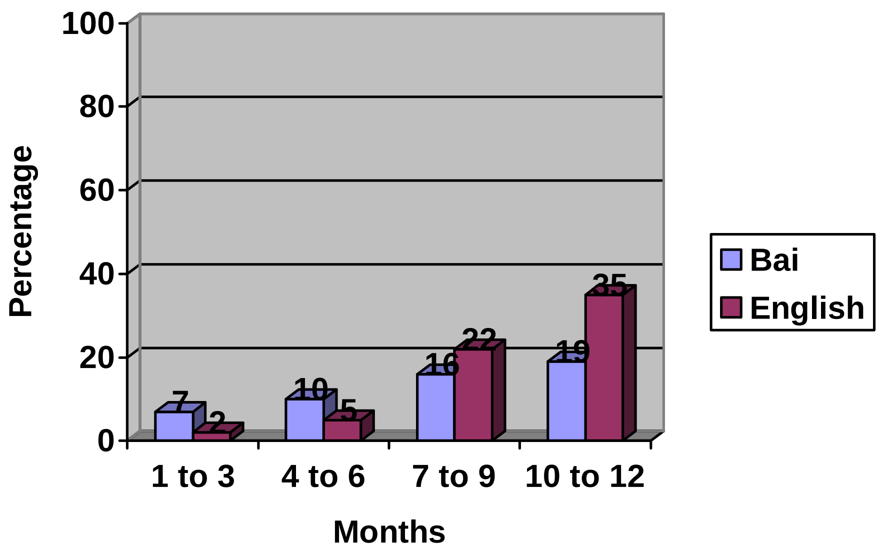
While the overall pattern of development of unconstricted utterances is the same for infants from both language groups, it is nonetheless noteworthy that the number of unconstricted utterances produced by the English infants rises more dramatically than is the case for the Bai infants, and that the overall number of unconstricted utterances is greater for English infants than for Bai infants. By the end of the first year, English infants produce nearly twice the number of unconstricted utterances that Bai infants produce. While unconstricted utterances comprise approximately one-fifth of Bai infants' production at the end of the first year, these utterances make up more than one-third of the English infants' utterances at that same time.

It would, however, be misleading to conclude that Bai infants are slower in developing the capacity to produce unconstricted utterances. While unconstricted utterances develop later than constricted utterances, the Bai infants produce more unconstricted utterances in months 1-3 and months 4-6 than the English infants. However, despite the comparatively early appearance of unconstricted utterances in the Bai infants' vocal repertoires, these vocalizations do not show a pattern of rapid acceleration, possibly demonstrating their relative unimportance or lack of salience in the target language, compared to other vocalization types.

**Table 4-3. Number of unconstricted utterances, English and Bai infants.**

| Age | Bai Infants | English Infants |
|---|---|---|
| Months 1-3 | 21 | 7 |
| Months 4-6 | 29 | 14 |
| Months 7-9 | 49 | 66 |
| Months 10-12 | 58 | 106 |
| Total | 136 | 193 |

**Figure 4-3. Percentage of unconstricted utterances, English and Bai infants.**




### *4.1.4 Summary*

The general pattern in the development of voice quality features for Bai and English infants shares many features in common, while exhibiting some differences that may reflect the influence of the target languages. These common and divergent patterns are demonstrated in Tables 4-4 and 4-5, and the corresponding Figures 4-4 and 4-5 for Bai and English infants, respectively. These tables and figures integrate the findings presented in Sections 4.1.1, 4.1.2, and 4.1.3.

**Table 4-4. Laryngeal quality in Bai utterances, by age.**

| Age | Constricted (n=734) | Dynamic (n=308) | Unconstricted (n=157) |
|---|---|---|---|
| 0-3 | 257 (86%) | 22 (7%) | 21 (7%) |
| 4-6 | 187 (62%) | 85 (28%) | 29 (10%) |
| 7-9 | 161 (54%) | 88 (29%) | 49 (16%) |
| 10-12 | 129 (43%) | 113 (38%) | 58 (19%) |

**Figure 4-4. Laryngeal quality in Bai utterances, by age and percentage.**

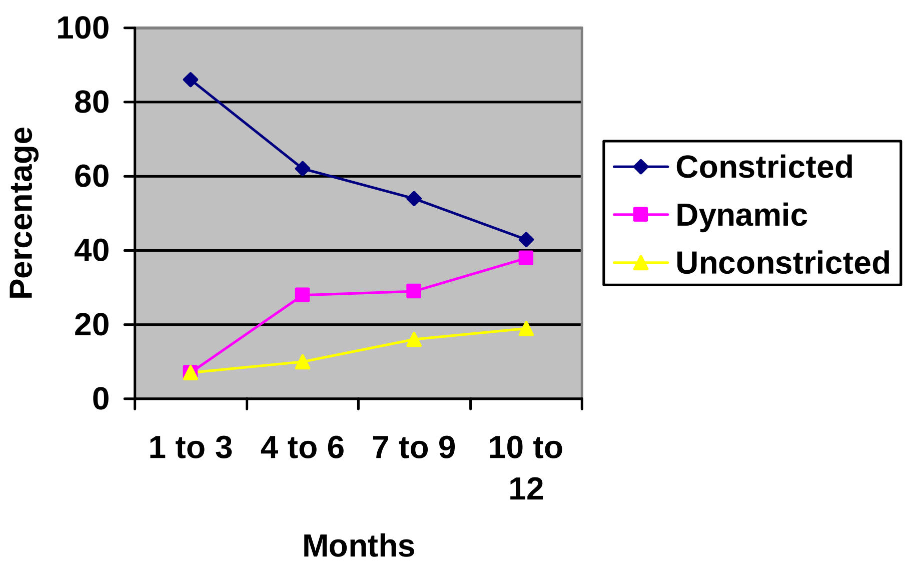
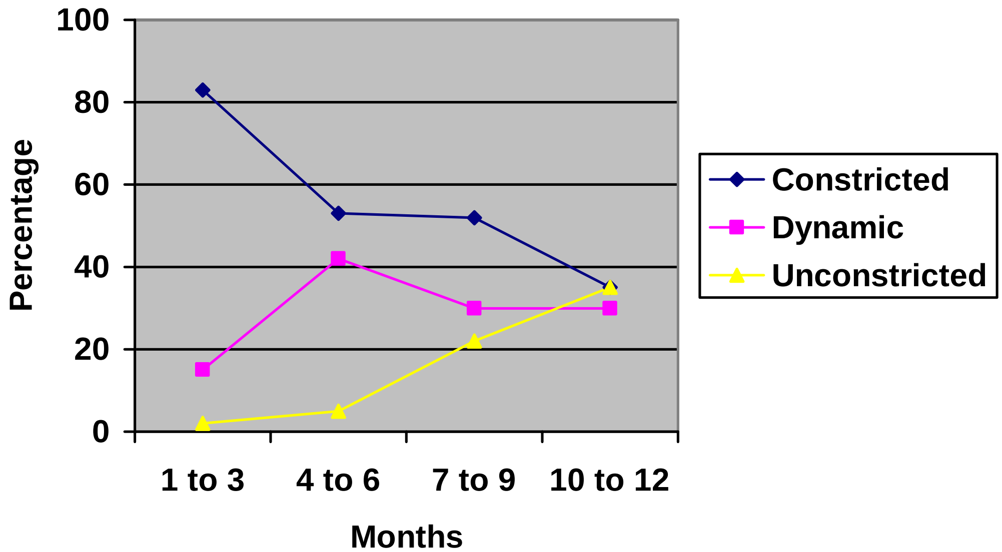



**Table 4-5. Laryngeal quality in English utterances, by age.**

| Age | Constricted (n=666) | Dynamic (n=350) | Unconstricted (n=193) |
|---|---|---|---|
| 0-3 | 248 (83%) | 45 (15%) | 7 (2%) |
| 4-6 | 158 (53%) | 127 (52%) | 14 (5%) |
| 7-9 | 155 (52%) | 89 (30%) | 66 (22%) |
| 10-12 | 105 (35%) | 89 (30%) | 106 (35%) |

**Figure 4-5. Laryngeal quality in English utterances, by age and percentage.**




For both language groups, laryngeal voice quality changes over time. Chi-square analysis of the relationship between laryngeal voice quality and age revealed a significant association between age and laryngeal voice quality for Bai infants ($\chi^2$ = 129.423, $p$ < .001) and English infants ($\chi^2$ = 239.015 , $p$ < .001). The strength of the association between laryngeal voice quality and age was moderate for both groups, though slightly weaker for the Bai infants (Cramer's V = .232), than for the English infants (Cramer's V = .314), most likely because of the slightly stronger persistence of constricted and dynamic utterances among the Bai infants, and the comparatively greater increase in unconstricted utterances among the English infants. Further chi-square analysis showed that differences in the development of laryngeal voice quality between the two language groups begin to become significant in months 10-12 ($\chi^2$ = 19.362 , $p$ < .001), though the strength of the association between language and laryngeal voice quality is not strong (Cramer’s V = .222).

Among infants from both language groups, constricted vocalizations are the most frequent utterances produced throughout every period of the first year of life, though they decline linearly as the year goes on, falling slightly faster and to a greater extent for English infants than for Bai infants. By contrast, unconstricted utterances gradually increase for infants from both language groups, though the increase is more dramatic for English infants than for Bai infants. In months 4-6, the number of dynamic utterances increases significantly in months 4-6 for infants of both language groups. Bai and English infants continue to produce a significant proportion of dynamic utterances throughout the first year. However, among the English infants, in the second half of the year, the proportion of dynamic utterances decreases relative to that produced in months 4-6, while for the Bai infants, dynamic utterances continue to increase during this period, with the highest incidence occurring in months 10-12.

The increasing proportion of dynamic utterances among the Bai infants, and the gradually decreasing proportion of such utterances among the English infants, in concert with the rise in unconstricted utterances in the latter group, may reflect the influence of the ambient language. In terms of the use of laryngeal voice quality features, Bai is a comparatively dynamic language that makes active use of laryngeal constriction and alternations between constricted and unconstricted voice quality settings. By contrast, English is a relatively static language in which the most typical laryngeal voice quality setting is unconstricted.

## 4.2 Utterance Types

This section outlines the general findings of this study on the development of syllabic and non-syllabic utterances, independent of laryngeal voice quality. In Section 4.2.1, I present and discuss the development of non-syllabic utterances in the first year of life. In Section 4.2.2, I describe changes in the incidence of mixed utterances throughout the first year of life. Section 4.2.3 discusses the development of syllabic utterances in the production of Bai and English infants. Finally, in Section 4.2.4, I summarize the findings on the development of utterance type categories for Bai and English infants, highlighting universal and language-specific patterns.

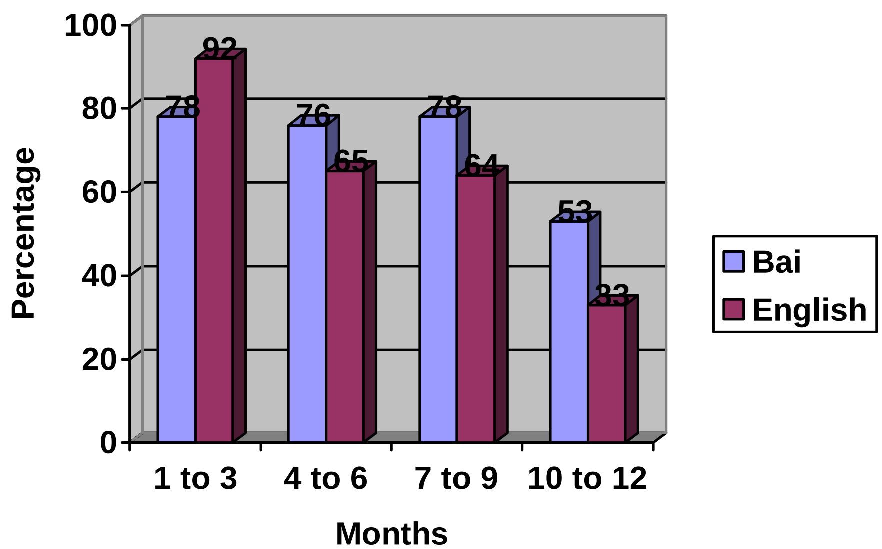
### *4.2.1 Non-Syllabic Utterances*

As shown in Table 4-6 and Figure 4-6, the overall pattern in the development of non-syllabic utterances is similar for Bai and English infants. Non-syllabic utterances predominate in months 1-3, and decline in frequency throughout the year for infants from both language groups. However, the rate and degree of decline differs between the two language groups. The Bai infants in the present study began to produce mixed and/or syllabic utterances before the English infants, as reflected in a 78% incidence of non-syllabic utterances for the Bai infants in months 1-3, versus a 92% incidence of such vocalizations for the English infants during this same period. However, the Bai infants' production of non-syllabic utterances remains relatively constant in months 4-6 and 7-9, before declining significantly in months 10-12. By contrast, English infants begin producing mixed and/or syllabic utterances in significant numbers only in months 4-6, but relative to the Bai infants, these utterances represent a greater proportion of their production for the remainder of the year. In months 10-12, only 33% of English infants' vocalizations are non-syllabic, compared to 53% for the Bai infants.

**Table 4-6. Number of non-syllabic utterances, by age and language.**

| Age | Bai (n=854) | English (n=763) |
|---|---|---|
| Months 1-3 | 234 | 277 |
| Months 4-6 | 227 | 194 |
| Months 7-9 | 234 | 192 |
| Months 10-12 | 159 | 100 |

**Figure 4-6. Percentage of non-syllabic utterances, by age and language.**




### *4.2.2 Mixed Utterances*

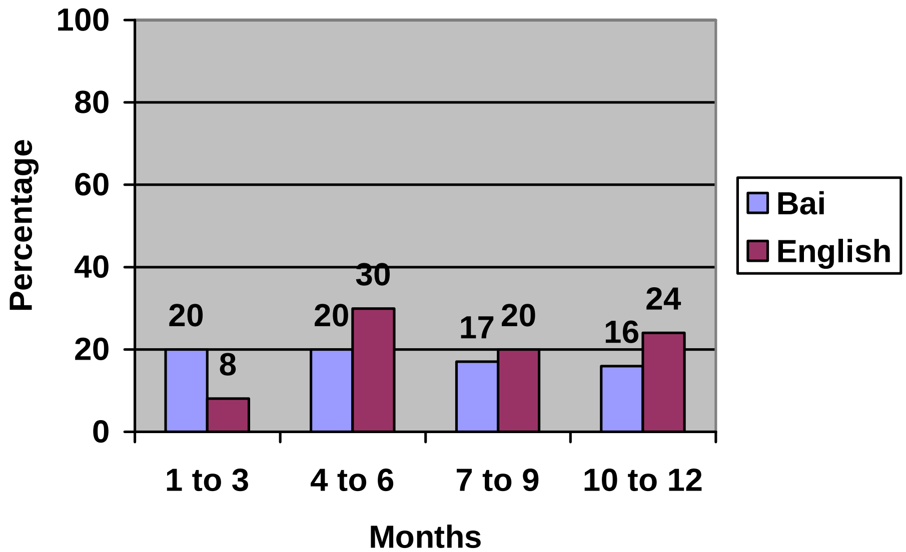
Within the theoretical framework adopted in the present study, much of speech development in the first year of life involves developing increasing control of, and integration of, the laryngeal and oral vocal tracts. Since laryngeal output is the substrate, or "carrier" of most oral sounds, some degree of laryngeal control necessarily precedes the emergence of oral control. Syllabic utterances, as defined in this study, represent a relatively advanced integration of the laryngeal and oral vocal tracts compared to non-syllabic utterances. However, syllabic utterances do not suddenly appear in infants' vocal repertoires, "from nowhere." Prior to their emergence, infants produce utterances that are primarily vocalic, but that involve a small degree of constriction in the oral vocal tract, such as that found in glides, approximants, and voiced dorsal fricatives. These constrictions are infants' first oral consonants and are the precursors to syllabic utterances.

Table 4-7 demonstrates that English and Bai infants begin to produce mixed utterances in months 1-3, and continue to produce such utterances throughout the remainder of the year on a steady basis. As illustrated in Figure 4-7 below, among the Bai infants, mixed utterances constitute a relatively constant proportion of their vocal repertoires throughout the year, comprising 20%, 20%, 17%, and 16% of their total production in months 1-3, 4-6, 7-9, and 10-12, respectively. Among English infants, mixed utterances begin as a comparatively small proportion of non-syllabic utterances in months 1-3, making up only 8% of their total production. However, mixed utterances increase dramatically in months 4-6, comprising 30% of their vocal output. In the remainder of the year, mixed utterances remain a strong, though declining, presence in their vocal repertoires, at 20% and 24% in months 7-9 and 10-12, respectively.

**Table 4-7. Number of mixed utterances, by age and language.**

| Age | Bai (n=217) | English (n=244) |
|---|---|---|
| Months 1-3 | 59 | 23 |
| Months 4-6 | 60 | 89 |
| Months 7-9 | 51 | 61 |
| Months 10-12 | 47 | 71 |

**Figure 4-7. Percentage of mixed utterances, by age and language.**




### *4.2.3 Syllabic Utterances*

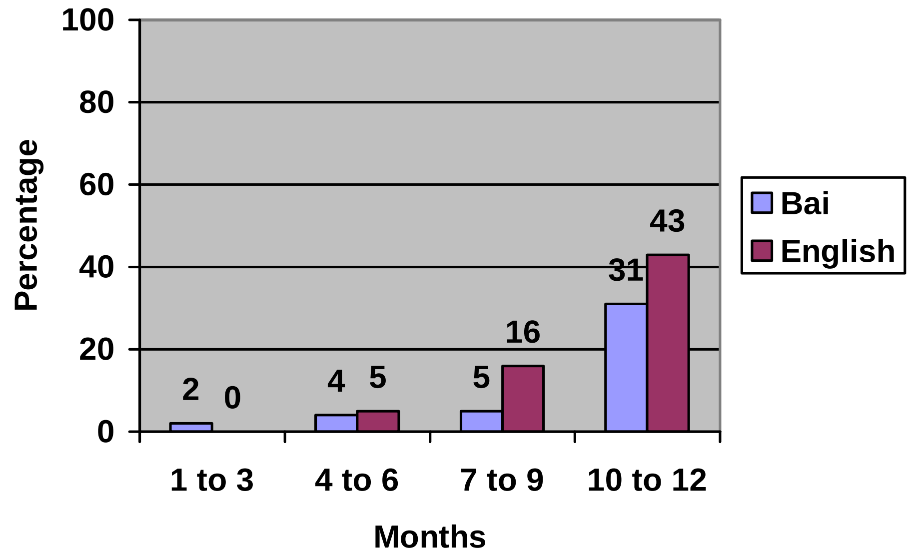
As shown in Table 4-8 and Figure 4-8 below, the two language groups differ in their production of syllabic utterances throughout the first year of life. Among the Bai infants, the proportion of syllabic utterances remains relatively constant throughout months 1-3, months 4-6, and months 7-9, before rising dramatically in months 10-12. By contrast, among the English infants, once syllabic utterances begin to appear in months 4-6, they steadily increase for the remainder of the year, tripling in proportion of total utterances produced during every successive quarterly period. Among the Bai infants, the production of syllabic utterances does not show rapid acceleration until the end of the first year.

The pattern in the incidence of syllabic utterances among the Bai versus English infants shows parallels with the pattern seen in the development of unconstricted utterances, discussed in Section 4.1.3 above. Just as unconstricted utterances develop later than constricted utterances, syllabic vocalizations develop later in the course of speech development than non-syllabic utterances for all infants. However, for the Bai infants, the development of syllabic utterances appears to be slower than for the English infants, despite the fact that these vocalizations appear in the Bai infants' vocal repertoires slightly earlier than in the English infants' production, albeit in small numbers. It is possible that the development of syllabic utterances presents greater complications for the Bai infants than for the English infants, given the interaction between laryngeal voice quality features and syllabic articulations in the target language. Specifically, Bai infants may need to acquire a different degree, or a different type, of laryngeal control prior to the full integration of laryngeal voicing with syllabic structures. Alternatively or in concert with this factor, it is possible that laryngeal voice quality features are at least as salient as syllabic structures in the ambient language of the Bai infants, and that they influence the infants' attention and output accordingly.

**Table 4-8. Number of syllabic utterances, by age and language.**

| Age | Bai (n=129) | English (n=190) |
|---|---|---|
| Months 1-3 | 7 | 0 |
| Months 4-6 | 13 | 14 |
| Months 7-9 | 15 | 47 |
| Months 10-12 | 94 | 129 |

**Figure 4-8. Percentage of syllabic utterances, by age and language.**




### *4.2.4 Summary*

As seen in the development of laryngeal voice quality, the patterns seen in the development of utterance type illustrate general and language-specific trends in speech development for infants from the two language groups studied. Chi-square analysis of the relationship between age and utterance type showed highly significant results for Bai infants ($\chi^2$ = 177.839, $p$ < .001) and for English infants ($\chi^2$ = 234.269, $p$ < .001). While the strength of the association between age and utterance type was strong for infants from both language groups, the association was slightly stronger for English infants (Cramer's V = .442) than for Bai infants (Cramer's V = .385), most likely because of the greater prevalence of non-syllabic utterances among the Bai infants, and the steadier development of syllabic utterances among the English infants. Further chi-square analysis of the relationship between age and utterance type showed that differences between the two language groups become significant in the second half of the year. In months 7-9, the two language groups differ significantly in the development of utterance type ($\chi^2$ = 18.419, $p$ < .001). While differences between the two groups remain significant in months 10-12, the differences are less marked ($\chi^2$ = 8.743, $p$ < .005). In both age periods, however, the strength of the association between language and utterance type is fairly weak, though in months 7-9, it is slightly stronger (Cramer's V = .175) than in months 10-12 (Cramer's V = .121).

In the first months of life, and throughout the first year, Bai and English infants produce mostly non-syllabic utterances. In the second half of the year, and particularly in months 10-12, syllabic utterances increase significantly in frequency. Prior to the emergence of syllabic vocalizations in large numbers, infants from both language groups begin to produce mixed vocalizations that involve the production of an oral constriction in what is primarily a vocalic utterance. Mixed utterances continue to comprise a substantial portion of non-syllabic vocalizations throughout the first year.

Despite these common trends, infants from the two language groups differ in striking ways, as shown in the summary Tables 4-9 and 4-10, and the associated Figures 4-9 and 4-10, for Bai and English infants, respectively.

**Table 4-9. Utterance types produced by Bai infants, by age.**

| Age | Non-Syllabic (n=854) | Mixed (n=217) | Syllabic (n=129) |
|---|---|---|---|
| Months 1-3 | 234 (78%) | 59 (20%) | 7 (2%) |
| Months 4-6 | 227 (76%) | 60 (20%) | 13 (4%) |
| Months 7-9 | 234 (78%) | 51 (17%) | 15 (5%) |
| Months 10-12 | 159 (53%) | 47 (16%) | 94 (31%) |

**Figure 4-9. Utterance types of Bai infants, by age and percentage.**

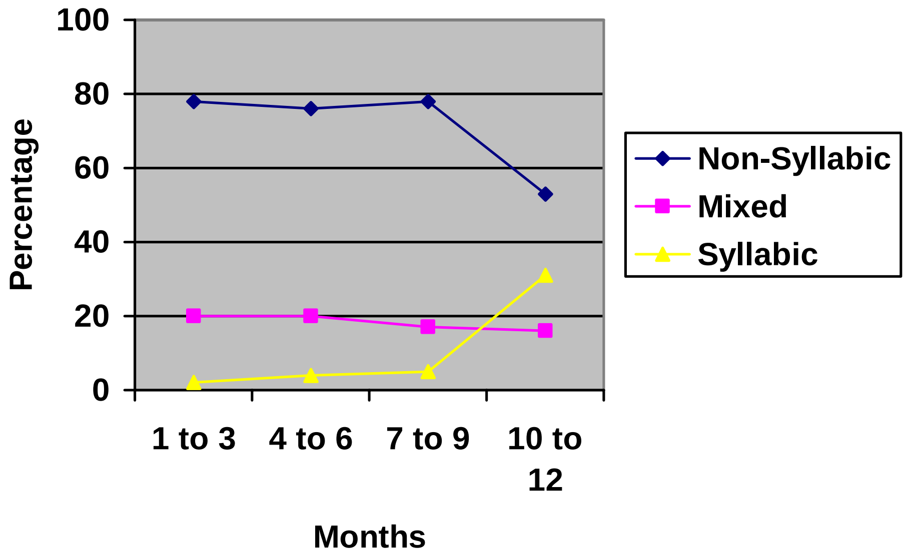
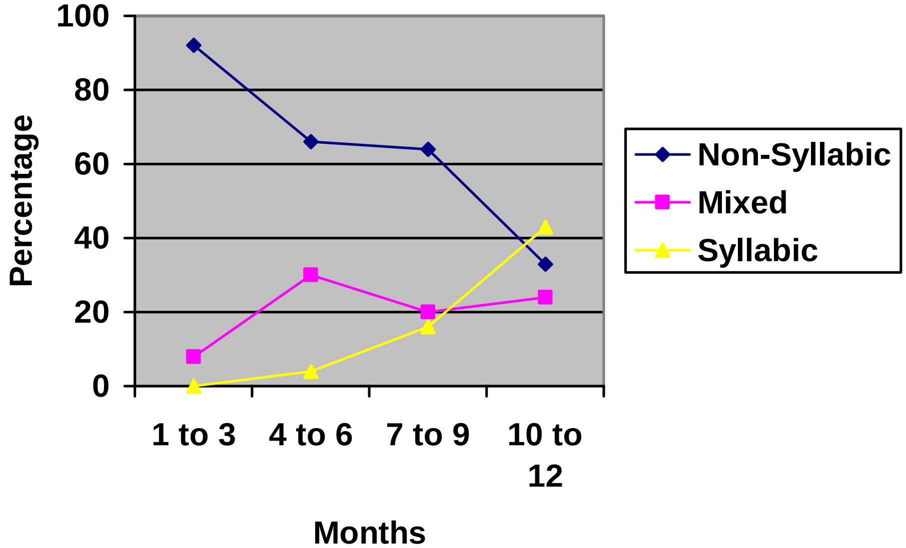



**Table 4-10. Utterance types produced by English infants, by age.**

| Age | Non-Syllabic (n=766) | Mixed (n=244) | Syllabic (n=190) |
|---|---|---|---|
| Months 1-3 | 277 (92%) | 23 (8%) | 0 (0%) |
| Months 4-6 | 197 (66%) | 89 (30%) | 14 (5%) |
| Months 7-9 | 192 (64%) | 61 (20%) | 47 (16%) |
| Months 10-12 | 100 (33%) | 71 (24%) | 129 (43%) |

**Figure 4-10. Utterance types of English infants, by age and percentage.**




Among Bai infants, syllabic utterances appear at least as early in speech development as they do among English infants, but they develop more slowly, only showing a significant increase in months 10-12. For English infants, once syllabic utterances begin to be produced, they proliferate. Moreover, with the exception of months 1-3, mixed utterances comprise a greater proportion of English infants' non-syllabic vocalizations throughout the first year. Relative to non-syllabic utterances, syllabic and mixed vocalizations involve a greater engagement of the oral vocal tract, and a larger degree of integration between the laryngeal and oral vocal tracts. Based on the patterns observed in this study, it appears that while infants from both language backgrounds are engaged in the same processes of phonetic development, these processes are structured differently between Bai and English infants. Once oral constriction and syllable structure begin to emerge, infants from both language groups are drawn to the further development of these phonetic capacities. However, based on the patterns observed in their vocal output, English infants are more attracted to the development of the oral vocal tract than are Bai infants, and achieve some degree of integration of the oral and laryngeal vocal tracts sooner than do Bai infants. These different patterns may reflect the relative importance of, and specific demands associated with, laryngeal control in English versus Bai. To explore this possibility, it is necessary to consider the relationship between laryngeal voice quality and utterance type—a topic to which we turn in Section 4.3 below.

## 4.3 The Relationship Between Laryngeal Voice Quality and Utterance Type

In this section, I will examine the distribution of laryngeal voice quality features in different utterance types. In Section 4.3.1, I consider the distribution of constricted, dynamic, and unconstricted voice quality in non-syllabic utterances. In Section 4.3.2, the incidence of these three voice quality features in mixed utterances is discussed. Section 4.3.3 focuses on the proportion of constricted, dynamic, and unconstricted voice quality in syllabic utterances. In Section 4.3.4, I summarize the results of this analysis.

### *4.3.1 Voice Quality in Non-Syllabic Utterances*

Tables 4-11 and Figure 4-11 illustrate the relative proportions of constricted, dynamic, and unconstricted vocalizations in the total number of non-syllabic vocalizations for months 1-3, 4-6, 7-9, and 10-12, for Bai infants. Table 4-12 and Figure 4-12 provide the same information for English infants.

Because non-syllabic utterances constitute the majority of utterances produced by both groups of infants for most of the first year, the patterns depicted here essentially replicate those already discussed in Section 4.1. For both language groups, constricted non-syllabic utterances decline linearly throughout the year, while unconstricted non-syllabic vocalizations steadily increase, with small differences in the degree and pace of increase between infants from the two language groups. Dynamic non-syllabic utterances gradually increase throughout the year for Bai infants, while among English infants, dynamic non-syllabic vocalizations increase dramatically in months 4-6, and then decrease slightly for the remainder of the year.

**Table 4-11. Laryngeal quality in Bai non-syllabic utterances.**

| Age | Constricted (n=579) | Dynamic (n=171) | Unconstricted (n=94) | Total (n=844) |
|---|---|---|---|---|
| 0-3 | 209 | 12 | 13 | 234 |
| 4-6 | 149 | 48 | 20 | 217 |
| 7-9 | 145 | 57 | 32 | 234 |
| 10-12 | 76 | 54 | 29 | 159 |

**Figure 4-11. Laryngeal quality in Bai non-syllabic utterances, by percentage.**

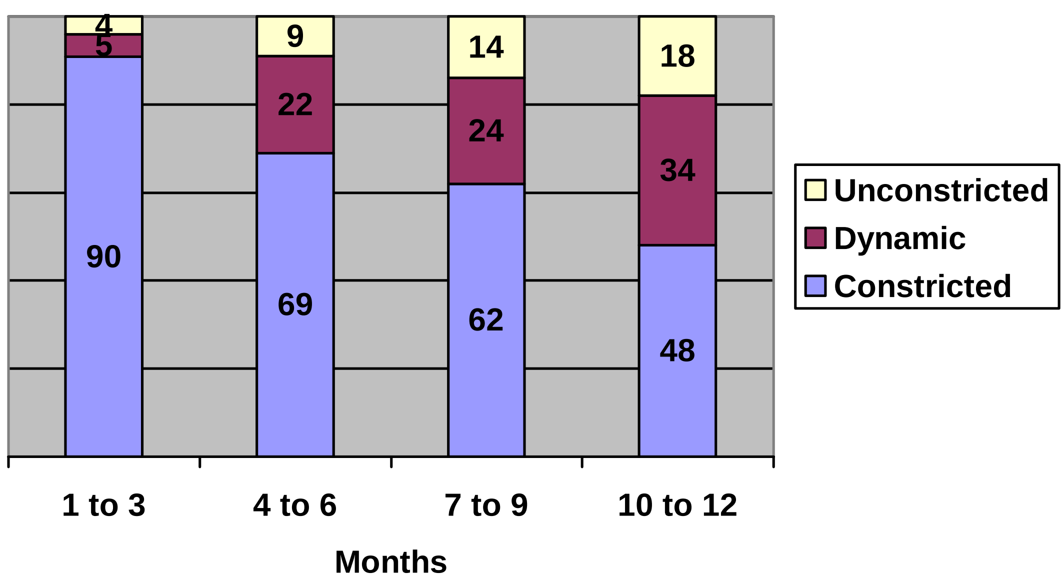
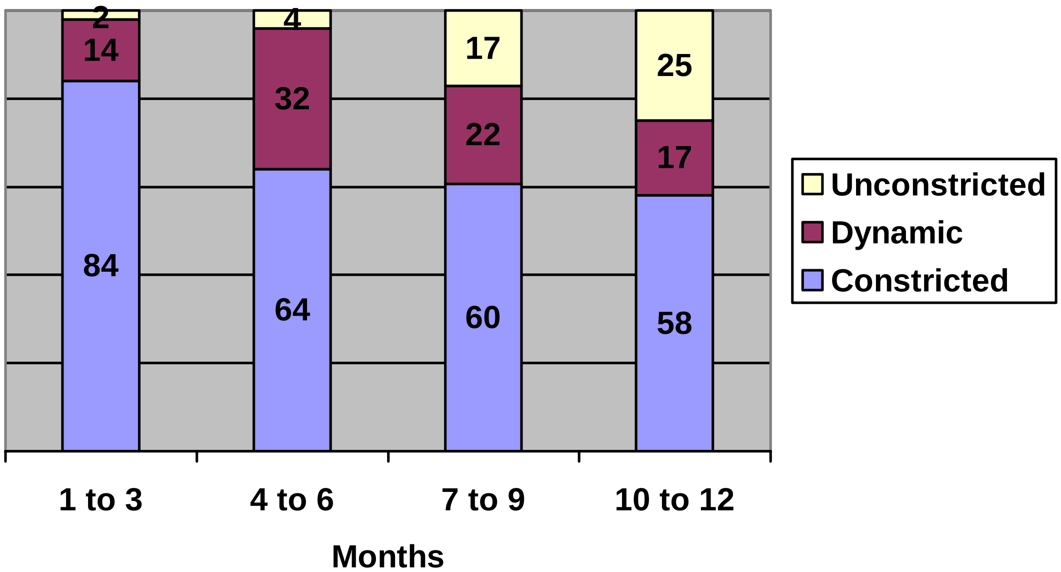


**Table 4-12. Laryngeal quality in English non-syllabic utterances.**

| Age | Constricted (n=531) | Dynamic (n=161) | Unconstricted (n=71) | Total (n=763) |
|---|---|---|---|---|
| 0-3 | 233 | 38 | 6 | 277 |
| 4-6 | 124 | 63 | 7 | 194 |
| 7-9 | 116 | 43 | 33 | 192 |
| 10-12 | 58 | 17 | 25 | 100 |

**Figure 4-12. Laryngeal quality in English non-syllabic utterances, by percentage.**




### *4.3.2 Voice Quality in Mixed Utterances*

Tables 4-13 and 4-14, and the corresponding Figures 4-13 and 4-14, show the distribution of laryngeal voice quality categories in the mixed vocalizations of Bai and English infants throughout the first year of life. In months 1-3, the majority of mixed utterances are constricted for infants from both language groups, demonstrating that in the first months of life, infants' first experiences with constriction in the oral vocal tract occur in the context of laryngeal constriction. For the remainder of the year, however, the majority of mixed utterances tend to be associated with dynamic laryngeal voice quality for Bai and English infants. Among both groups of infants, mixed utterances are least likely to be unconstricted, although unconstricted mixed utterances are more common for Bai and English infants in months 7-9 than they are in months 1-3, 4-6, and 10-12. Also, English infants produce more unconstricted mixed utterances in months 7-9 and 10-12 than do Bai infants, though the numbers may be too small to be meaningful. Overall, throughout the year, mixed utterances are most likely to be dynamic, and least likely to be unconstricted.

**Table 4-13. Laryngeal quality in Bai mixed utterances.**

| Age | Constricted (n=105) | Dynamic (n=84) | Unconstricted (n=28) | Total (n=217) |
|---|---|---|---|---|
| 1-3 | 46 | 8 | 5 | 59 |
| 4-6 | 26 | 29 | 5 | 60 |
| 7-9 | 14 | 25 | 12 | 51 |
| 10-12 | 19 | 22 | 6 | 47 |

**Figure 4-13. Laryngeal quality in Bai mixed utterances, by percentage.**

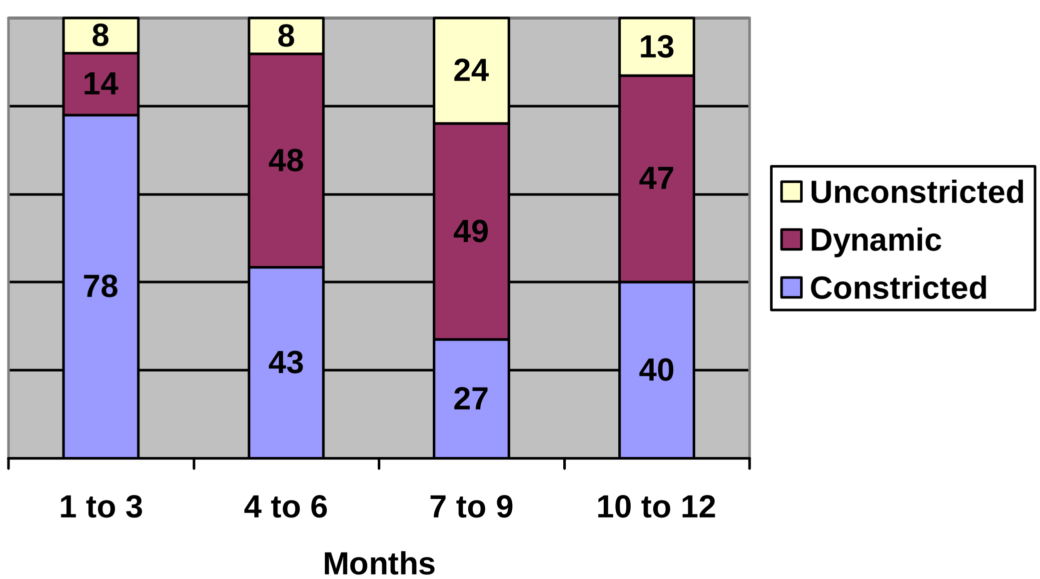
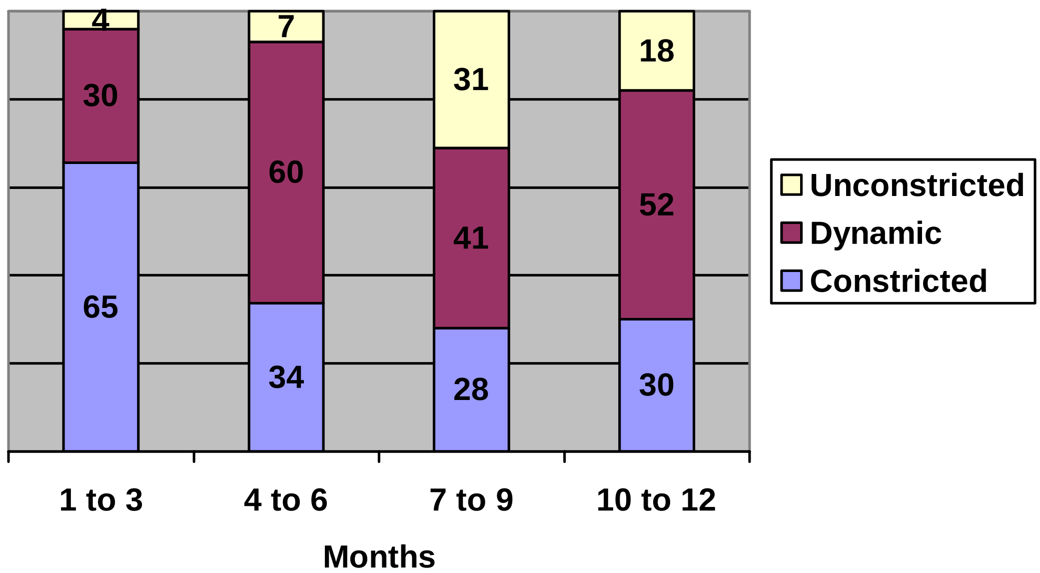


**Table 4-14. Laryngeal quality in English mixed utterances.**

| Age | Constricted (n=83) | Dynamic (n=122) | Unconstricted (n=39) | Total (n=244) |
|---|---|---|---|---|
| 1-3 | 15 | 7 | 1 | 23 |
| 4-6 | 30 | 53 | 6 | 89 |
| 7-9 | 17 | 25 | 19 | 61 |
| 10-12 | 21 | 37 | 13 | 71 |

**Figure 4-14. Laryngeal quality in English mixed utterances, by percentage.**




Mixed dynamic utterances may play a special role in the process of speech development for infants from both language groups. Within the context of this study, dynamic utterances are those most likely to be associated with "vocal play," and thus, with infants' exploration of their vocal ranges in terms of pitch and constriction. As such, dynamic utterances are an important medium for the development of laryngeal control for infants of both language backgrounds. In turn, within this study, mixed utterances are understood as the context in which oral consonants first begin to emerge, and as intermediary between non-syllabic and syllabic utterances. The fact that mixed utterances are most likely to be dynamic in laryngeal voice quality throughout the first year suggests that vocal play provides a medium in which infants not only develop laryngeal voice quality, but in which infants are most likely to begin to integrate their emergent control of the oral vocal tract with their developing control of the laryngeal vocal tract.

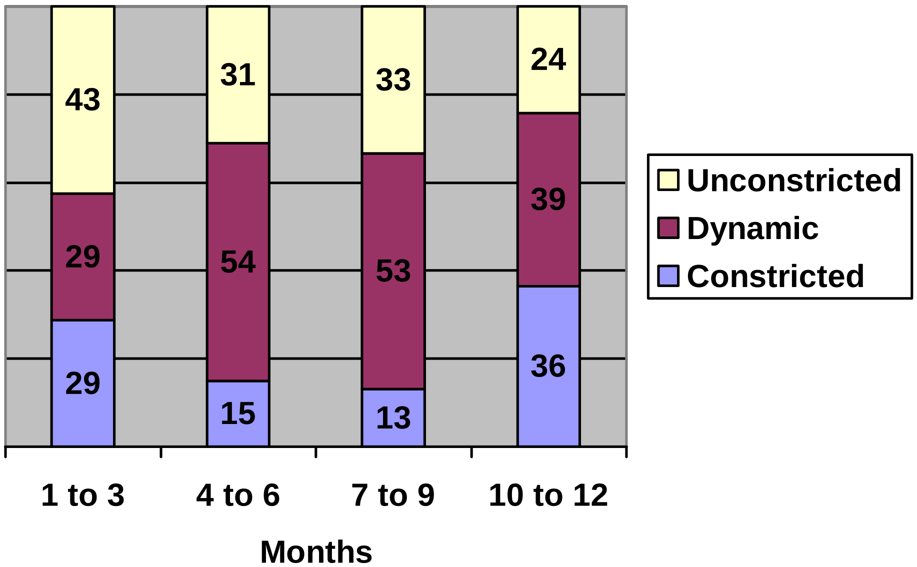
### *4.3.3 Voice Quality in Syllabic Utterances*

Table 4-15 and Figure 4-15 show the distribution of constricted, dynamic, and unconstricted voice quality categories in syllabic utterances for Bai infants in months 1-3, 4-6, 7-9, and 10-12. As discussed in Section 4.2 above, Bai infants do not begin to produce syllabic utterances in significant numbers until months 10-12, making comparisons between voice quality settings throughout the year difficult. However, it is noteworthy that in months 10-12, the syllabic utterances of Bai infants are distributed relatively evenly across laryngeal voice quality categories, at 36%, 39%, and 24% for constricted, dynamic, and unconstricted categories, respectively. This distribution contrasts not only with the distribution of laryngeal voice quality among English syllabic utterances, to be discussed below, but with the distribution of laryngeal voice quality in the Bai non-syllabic utterances for the same period, which is 46%, 37%, and 17% for constricted, dynamic, and unconstricted non-syllabic utterances, respectively.

**Table 4-15. Laryngeal quality in Bai syllabic utterances.**

| Age | Constricted (n=40) | Dynamic (n=54) | Unconstricted (n=35) | Total |
|---|---|---|---|---|
| 1-3 | 2 | 2 | 3 | 7 |
| 4-6 | 2 | 7 | 4 | 13 |
| 7-9 | 2 | 8 | 5 | 15 |
| 10-12 | 34 | 37 | 23 | 94 |

**Figure 4-15. Laryngeal quality in Bai syllabic utterances, by percentage.**

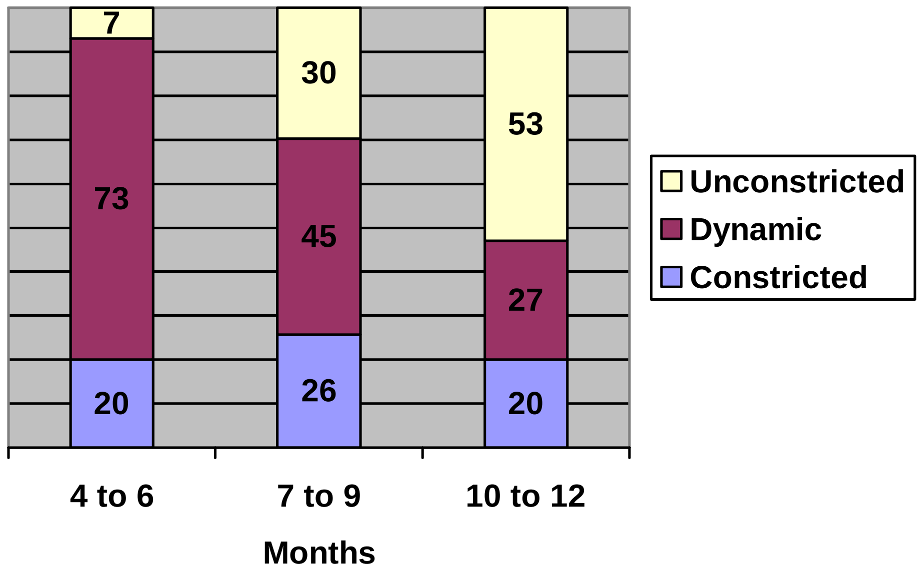



The distribution of laryngeal voice quality in the syllabic utterances of Bai infants contrasts markedly with that seen among the English infants, shown in Table 4-16 and Figure 4-16 below.

**Table 4-16. Laryngeal quality in English syllabic utterances.**

| Age | Constricted (n=41) | Dynamic (n=67) | Unconstricted (n=83) | Total |
|---|---|---|---|---|
| 0-3 | 0 | 0 | 0 | 0 |
| 4-6 | 3 | 11 | 1 | 15 |
| 7-9 | 12 | 21 | 14 | 47 |
| 10-12 | 26 | 35 | 68 | 129 |

**Figure 4-16. Laryngeal quality in English syllabic utterances, by percentage.**




The contrast between the English and Bai infants is especially striking in months 10-12, the only period for which a meaningful comparison between the two language groups can be made, given the relative lack of syllabic utterances among the Bai infants in months 7-9. In months 10-12, 53% of syllabic utterances produced by English infants are unconstricted, compared to 24% among the Bai infants. While the proportion of unconstricted utterances increases significantly among English infants for syllabic and non-syllabic utterances, the increase in unconstricted utterances is more dramatic for syllabic than for non-syllabic utterances. For English infants, the proportion of non-syllabic unconstricted utterances stands at 4%, 17%, and 25% in months 4-6, 7-9, and 10-12, respectively. By comparison, in months 4-6, 7-9, and 10-12, the only periods for which comparisons are possible, the incidence of unconstricted syllabic utterances is 7%, 30%, and 53%, respectively.

In terms of laryngeal voice quality, the gulf between syllabic and non-syllabic utterances widens progressively as the year goes on, and is more marked than seen for the

Bai infants. These differences likely reflect the influence of the ambient language. Syllabic utterances are the most speech-like of the utterances considered in this study. Bai infants are learning to speak a language in which laryngeal constriction, and alternations between constricted and unconstricted settings, occur at the syllabic level, a pattern that seems to be reflected in the distribution of laryngeal voice quality settings in months 10-12. By contrast, English infants are learning to speak a language in which voice quality can range according to a variety of individual and contextual factors, but in which most speech is unconstricted.

### *4.3.4 Summary*

In Section 4.3, I have presented the findings on the distribution of laryngeal voice quality in non-syllabic, mixed, and syllabic utterances. As discussed in Sections 4.3.1 and 4.3.2, non-syllabic utterances are most likely to be constricted for infants from both language groups throughout the year, while mixed utterances are most likely to be dynamic throughout the year. Syllabic utterances only begin to appear in significant numbers for both language groups in months 10-12, by which time, the distribution of laryngeal voice quality begins to show significant differences. Chi-square analysis of the relationship between utterance type and laryngeal voice quality by language showed no significant language differences for non-syllabic and mixed utterances ($\chi^2 = 1.612$, $p = .204$), but highly significant differences for syllabic utterances ($\chi^2 = 14.199$, $p < .001$). However, for syllabic utterances, the strength of the association between laryngeal voice quality and utterance type by language was fairly weak (Cramer's V = .222).

## 4.4 Overview

In this final section of Chapter 4, I summarize the universal and language-specific patterns identified in the present study of infant vocalizations produced by Bai and English infants in the first year of life. Table 4-17 below lists the patterns that are shared between the Bai and English infants for the laryngeal voice quality and utterance type categories considered in this study. The most clearly documented universal trends are the decrease in production of constricted utterances and the increase in production of unconstricted utterances throughout the year, along with the gradual decrease in non-syllabic utterances and the significant increase in syllabic vocalizations in months 10-12. Also, infants from both language groups produce a significant proportion of dynamic utterances throughout the year; among all the utterances considered in this study, dynamic utterances are the ones most representative of "vocal play." Finally, infants produce mixed utterances in significant numbers throughout the year. For infants from both language groups, mixed utterances first appear in months 1-3, and are most commonly produced with laryngeal constriction at that time. For the balance of the year, however, mixed utterances are most commonly produced with dynamic alternations between constricted and unconstricted settings, and thus, are most often produced in the context of "vocal play." It is in the context of mixed utterances that infants' first oral consonants appear. Thus, mixed utterances represent a degree of integration of the laryngeal and oral vocal tracts.

**Table 4-17. Universal production patterns by laryngeal quality and utterance type.**

| Category | Universal Pattern |
|---|---|
| Constricted Utterances | • Decrease in frequency throughout the first year. |
| Dynamic Utterances | • Increase significantly in months 4-6 and comprise a notable proportion of total utterances for the balance of the first year. |
| Unconstricted Utterances | • Increase in frequency throughout the first year. |
| Non-Syllabic Utterances | • Decrease in frequency throughout the first year.<br>• Most commonly produced with laryngeal constriction throughout the year, though the incidence of dynamic and unconstricted non-syllabic utterances steadily increases. |
| Mixed Utterances | • Begin to appear early in development, in months 1-3, and decline somewhat in frequency in the second half of the first year.<br>• Occur most commonly in constricted settings in months 1-3, but then most frequently in dynamic settings for the balance of the first year. |
| Syllabic Utterances | • Increase in frequency throughout the first year, with the greatest increase in months 10-12.<br>• No striking universal pattern in terms of laryngeal voice quality. |

Tables 4-18 and 4-19 below list the patterns observed in this study that are unique to the Bai and English infants in the sample, respectively. Bai and English infants differ in the degree to which constricted utterances decrease in frequency, and in the extent to which unconstricted utterances increase in frequency, throughout the year. Among the English infants, the decrease in constricted utterances, and the increase in unconstricted utterances, is more marked than among the Bai infants. Also, syllabic utterances increase more quickly, and to a greater extent, for the English infants than for the Bai infants. While syllabic utterances are considered the most "speech-like" of the utterances examined in this study, the division between syllabic and non-syllabic may be less distinct for the Bai infants than for the English infants. Given the importance of pitch-dependent voice quality variations in Bai, it is possible that the sort of vocal practice that the Bai infants seem to be engaging in to a greater degree than the English infants reflects the greater salience of pitch and voice quality variations in Bai. Finally, the laryngeal voice quality settings in the syllabic utterances differs strikingly between the two language groups in months 10-12. By this time, most English syllabic utterances are unconstricted, while Bai syllabic utterances are relatively evenly distributed among constricted, dynamic, and unconstricted settings, reflecting the distribution of laryngeal voice quality features in the target language of each group of infants. In Chapter 6, these findings will be further discussed in relation to the literature on infant speech reviewed in Chapter 2, and in relation to the results of the adult perceptual test, the results of which are reported in Chapter 5.

**Table 4-18. Bai production patterns by laryngeal quality and utterance type.**

| Category | Bai-specific Pattern |
|---|---|
| Constricted Utterances | • Decline less quickly and to a lesser degree than seen among English infants. |
| Dynamic Utterances | • Increase significantly in months 4-6 and continue to increase slightly in frequency for the balance of the first year. |
| Unconstricted Utterances | • Increase slowly in frequency throughout the year. |

| | |
|---|---|
| Non-Syllabic Utterances | • Steadily produced through to the end of month 9, and then decrease significantly in months 10-12.<br>• Most commonly produced with laryngeal constriction, but by months 10-12, more than half of non-syllabic utterances are dynamic or unconstricted, |
| Mixed Utterances | • No striking language-specific tendencies. |
| Syllabic Utterances | • Appear early in development, but do not increase significantly in frequency till months 10-12.<br>• By months 10-12, fairly evenly distributed among constricted, dynamic, and unconstricted laryngeal settings, compared to English syllabic utterances. |

**Table 4-19. English production patterns by laryngeal quality and utterance type.**

| Category | English-specific Pattern |
|---|---|
| Constricted Utterances | • Decline more quickly and to a greater extent than seen among Bai infants. |
| Dynamic Utterances | • Increase dramatically in months 4-6 and then decrease somewhat in frequency for the balance of the first year. |
| Unconstricted Utterances | • Increase dramatically in frequency throughout the first year. |
| Non-Syllabic Utterances | • Decrease significantly in months 4-6 and then again dramatically in months 10-12.<br>• Most commonly produced with laryngeal constriction throughout the year. |
| Mixed Utterances | • No striking language-specific tendencies. |
| Syllabic Utterances | • Proliferate throughout the first year once they are initially produced in months 4-6.<br>• By months 10-12, most syllabic utterances are unconstricted. |

# Chapter 5

## PERCEPTION RESULTS

In this chapter, I present and discuss the findings of the second part of this study, which concerns the perception of infant vocalizations by English and Bai adult listeners. As outlined in Chapter 3, in the perceptual task, 40 English and 40 Bai participants listened to 36 infant vocalizations, which were presented in a random order. Participants were asked to rate each sound on a scale of one to five, based on their judgment of how important the sound was in learning to speak their native language (English or Bai). Following the listening task, participants were interviewed for approximately five minutes and were asked to reflect on the factors that influenced their ratings of the infant vocalizations.

Section 5.1 outlines the Bai and English ratings for the laryngeal voice quality categories constricted, dynamic, and unconstricted, including mean ratings for both genders between and within languages. Section 5.2 provides a similar analysis for the utterance categories non-syllabic and syllabic. In Section 5.3, I provide the results of the listening task for the subcategories constricted non-syllabic, dynamic non-syllabic, and unconstricted non-syllabic, while in Section 5.4, results for the sub-categories constricted syllabic, dynamic syllabic, and unconstricted syllabic are presented. In Sections 5.3 and 5.4, I identify universal and language-specific preferences, based on shared and divergent patterns in the ratings from both language groups within each sub-category. I also discuss gender differences in perceptual preferences within each language group. Section 5.5 outlines the findings of the qualitative component of the perceptual study. Finally, Section 5.6 summarizes the results of the perception study.

## 5.1 Laryngeal Voice Quality

In this section, I present the findings from the listening task for the categories constricted, dynamic, and unconstricted, independent of utterance type, by language and gender. This analysis highlights general trends in listeners' perceptions of the importance of laryngeal voice quality variations in language acquisition. Except for between-language comparisons, statistical analyses of the results are reported separately for each language, because Levene's test of homogeneity of variance showed significant differences in variance between the two language groups, but no significant differences in variance within language groups.

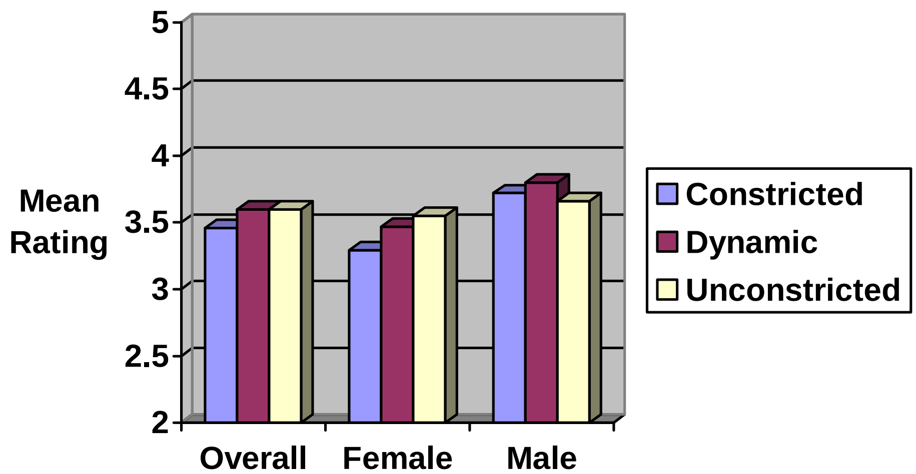
### *5.1.1 Bai Ratings*

As shown in Table 5-1 and Figure 5-1 below, overall, Bai listeners do not show an obvious preference for any of the laryngeal voice quality categories featured in the listening task. Statistical analysis showed that constriction had no significant effect on Bai listeners' ratings ($F(2, 476) = 2.436$, $p < .1$, $r = .006$). Mean ratings for constricted, dynamic, and unconstricted utterances are 3.46, 3.60, and 3.60, respectively. However, women and men differ slightly, though not significantly, in their ratings of laryngeal voice quality ($F(2, 476) = 2.646$, $p < .1$, $r = .007$). Female Bai listeners' mean ratings for constricted, dynamic, and unconstricted utterances were 3.29, 3.47, and 3.55, respectively. Among female listeners, constricted utterances received the lowest average ratings, and unconstricted utterances the highest, with dynamic utterances falling in between. By contrast, male Bai listeners rated dynamic utterances the highest, and unconstricted utterances the lowest, with constricted utterances falling in between. However, it is important to note that overall, as well as between and within genders, the differences in mean ratings for these utterance categories is slight, and responses range widely. The only striking difference in the mean ratings of laryngeal voice quality is that between male and female ratings of constricted utterances, which are 3.72 and 3.29, respectively. Compared to Bai males, Bai females seem to exhibit a slight aversion to constricted utterances. However, it is also true that mean Bai male ratings are higher than mean Bai female ratings for every laryngeal voice quality, so some of the differences in the category ratings may reflect differences in the ways that women and men carried out the task, rather than an inherent preference or aversion to any laryngeal voice quality category. It is necessary to examine ratings of sub-categories, and individual ratings within sub-categories, to determine the possible source of the observed difference.

**Table 5-1. Bai ratings of laryngeal quality, overall and by gender.**

| Laryngeal Category | Overall Mean (s.d.) (n=40) | Female Mean (s.d.) (n=24) | Male Mean (s.d.) (n=16) |
|---|---|---|---|
| Constricted | 3.46 (1.52) | 3.29 (1.53) | 3.72 (1.47) |
| Dynamic | 3.60 (1.42) | 3.47 (1.37) | 3.80 (1.47) |
| Unconstricted | 3.60 (1.44) | 3.55 (1.42) | 3.66 (1.48) |

**Figure 5-1. Bai mean ratings of laryngeal quality, overall and by gender.**



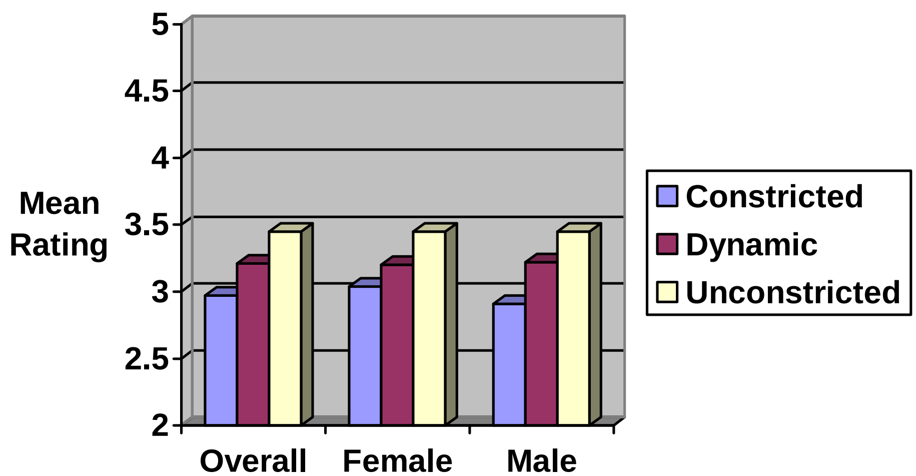

### *5.1.2 English Ratings*

In contrast to the Bai listeners, laryngeal voice quality played a significant role in influencing English listeners' ratings ($F(2, 476) = 27.941$, $p < 0.001$, $r = .205$). Mean English ratings for constricted, dynamic, and unconstricted utterances were 2.97, 3.21, and 3.45, respectively. As shown in Table 5-2 and Figure 5-2, there was virtually no difference in mean ratings for laryngeal voice quality categories between genders ($F(2, 476) = 1.559$, $p < 0.1$, $r = .014$). Mean ratings for women and men are almost identical, with the exception of the mean ratings for constricted utterances. Men rated constricted utterances slightly lower than women, with a mean rating of 2.91 for the men, versus 3.04 for the women. These findings suggest that, as a group, English listeners prefer unconstricted utterances over dynamic utterances, and dynamic utterances over constricted utterances, echoing the pattern seen among the Bai females, though the pattern is clearer among the English listeners. As with the Bai listeners, responses ranged widely, as seen in the standard deviations reported in Table 5-2. Ratings for utterance type, combined with the findings of the qualitative component of this study, account for much of the variation in the ratings.

**Table 5-2. English ratings of laryngeal quality, overall and by gender.**

| Laryngeal Category | Overall Mean (s.d.) (n=40) | Female Mean (s.d.) (n=19) | Male Mean (s.d.) (n=21) |
|---|---|---|---|
| Constricted | 2.97 (1.35) | 3.04 (1.38) | 2.91 (1.33) |
| Dynamic | 3.21 (1.29) | 3.20 (1.34) | 3.22 (1.24) |
| Unconstricted | 3.45 (1.28) | 3.45 (1.27) | 3.45 (1.30) |

**Figure 5-2. English mean ratings of laryngeal voice quality, overall and by gender.**



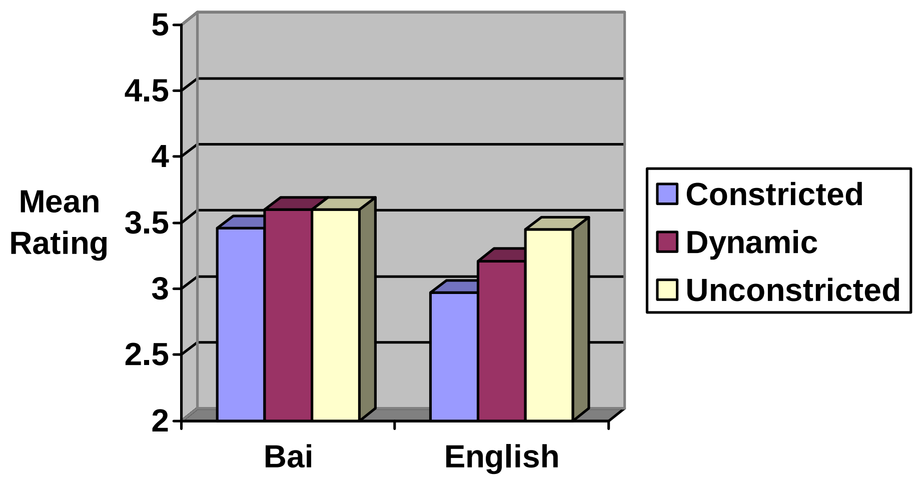

### *5.1.3 Cross-Linguistic Comparison*

The cross-linguistic comparison of mean ratings for laryngeal type highlights the fact that while Bai listeners show no obvious preference for any of the three laryngeal categories in this study, English listeners show a consistent pattern of rating constricted utterances lower than dynamic utterances, and dynamic utterances lower than unconstricted utterances. Statistical analysis showed a significant difference in laryngeal voice quality ratings by language (*F*(2, 952) = 8.290, $p > .05$, $r = .013$). The overall mean ratings for constricted, dynamic, and unconstricted utterances for each language, reported in Sections 5.1.1 and 5.1.2 above, are repeated for convenience in Table 5-3 below, and illustrated in Figure 5-3. As already noted, ratings for both language groups range widely. However, as seen in Table 5-3, standard deviations for Bai listeners tend to be larger than for English listeners, suggesting that the rating task may be affected by different factors for listeners from the two language groups and/or that as a group, English listeners are more homogeneous in their perceptions of laryngeal voice quality than are Bai listeners.

**Table 5-3. Bai and English ratings of laryngeal quality, overall.**

| Laryngeal Category | Bai Mean (s.d.) (n=40) | English Mean (s.d.) (n=40) |
|---|---|---|
| Constricted | 3.46 (1.52) | 2.97 (1.35) |
| Dynamic | 3.60 (1.42) | 3.21 (1.29) |
| Unconstricted | 3.60 (1.44) | 3.45 (1.28) |

**Figure 5-3. Mean ratings of laryngeal quality, by language.**



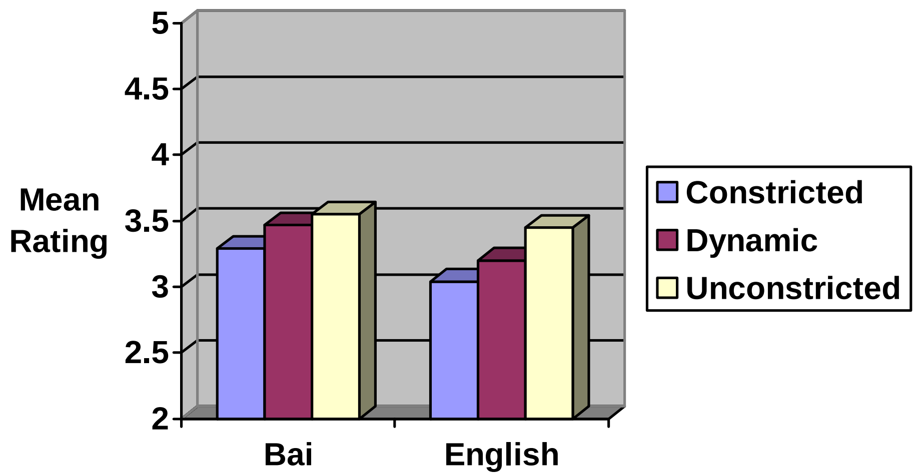

*5.1.3.1 Female Ratings*

In this section, ratings of female listeners from each language group are compared. Table 5-4 and Figure 5-4 below show that Bai and English women exhibit the same relative preferences with respect to laryngeal voice quality. Among both groups of women, constricted utterances receive the lowest mean ratings and unconstricted utterances the highest, with mean ratings for dynamic utterances falling in between. Given that women from both groups make very different uses of laryngeal constriction in their respective languages, it is likely that the similarities in the patterns of mean ratings reflect extra-linguistic, rather than specifically linguistic, factors. It should also be noted that since, as discussed in Section 5.1.2, English female ratings were highly similar to English male ratings, the Bai women's mean ratings for laryngeal voice quality are similar not only to those of English women, but also to those of English men.

**Table 5-4. Female ratings of laryngeal quality, by language.**

| Laryngeal Category | Bai Female Mean (s.d.) (n=24) | English Female Mean (s.d.) (n=19) |
|---|---|---|
| Constricted | 3.29 (1.53) | 3.04 (1.38) |
| Dynamic | 3.47 (1.37) | 3.20 (1.34) |
| Unconstricted | 3.55 (1.42) | 3.45 (1.27) |

**Figure 5-4. Female mean ratings of laryngeal quality, by language.**



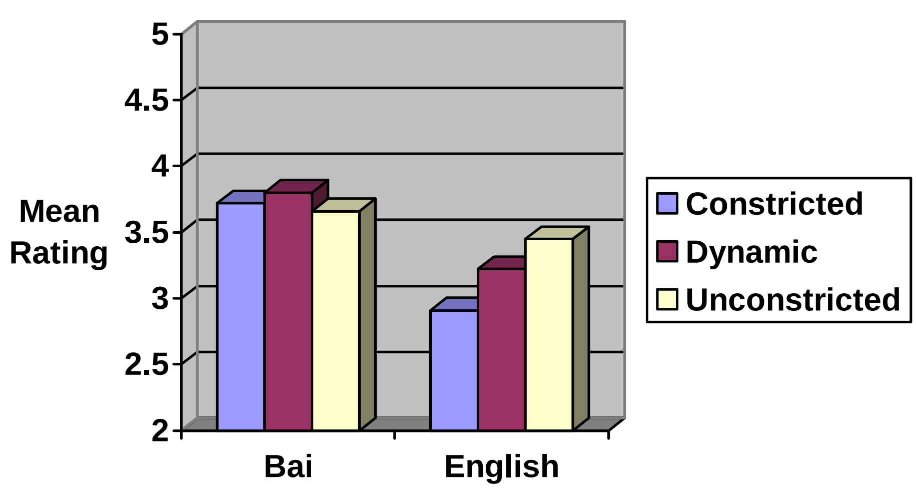

*5.1.3.2 Male Ratings*

As detailed in Table 5-5 and illustrated in Figure 5-5 below, Bai and English males show a marked contrast in their mean ratings of laryngeal voice quality. While men from both language groups have similar mean ratings for unconstricted utterances, Bai males rate constricted and dynamic utterances considerably higher than do English males. The most striking contrast lies in the mean rating for constricted utterances, which is 3.72 for Bai males and 2.91 for English males. Given the different use of laryngeal constriction in Bai and English, it is tempting to consider that the differences in the Bai and English male ratings of laryngeal constriction stem from linguistic factors. However, strictly speaking, if this were the case, one would expect the Bai female ratings of laryngeal quality to be similar to those of the Bai male ratings. However, as shown in Section 5.1.3.1 above, the Bai female ratings of laryngeal quality exhibit a pattern similar to that seen among English males and females, though the differences in mean ratings between laryngeal categories is less marked than that seen among the English listeners. In the latter respect, the Bai female listeners are more similar to the Bai male listeners than to the English listeners. To clarify the relative influence of linguistic and extra-linguistic factors in these ratings, it is necessary to delve further into the data.

**Table 5-5. Male ratings of laryngeal quality, by language.**

| Laryngeal Category | Bai Male Mean (s.d.) (n=16) | English Male Mean (s.d.) (n=21) |
|---|---|---|
| Constricted | 3.72 (1.47) | 2.91 (1.33) |
| Dynamic | 3.80 (1.47) | 3.22 (1.24) |
| Unconstricted | 3.66 (1.48) | 3.45 (1.30) |

**Figure 5-5. Male mean ratings of laryngeal quality, by language.**



### *5.1.4 Summary*

In Section 5.1, I have presented the mean ratings for constricted, dynamic, and unconstricted utterances, by language and gender. Overall, the findings suggest that laryngeal quality, as a unitary factor, operates more strongly in influencing the ratings of the English listeners than it does the Bai listeners. While differences in the relative rankings of constricted, dynamic, and unconstricted utterances are not large, they are clear and highly consistent between genders. The perception of laryngeal constriction in infant vocalizations is more complicated among the Bai listeners. While female Bai listeners' ratings show a pattern that is similar to that found for English listeners, the pattern is distributed over a narrower range, suggesting a less distinct division between the laryngeal categories used to structure the listening task in the present study. This lack of distinction is most clearly seen in the mean ratings of male Bai listeners, whose ratings of constricted, dynamic, and unconstricted utterances differ minimally.

## 5.2 Utterance Type

In this section, I present the mean average ratings for non-syllabic and syllabic utterances, independent of laryngeal voice quality, highlighting universal and language-specific patterns in the participants' ratings, and, where relevant, gender-based differences in ratings of utterance type.

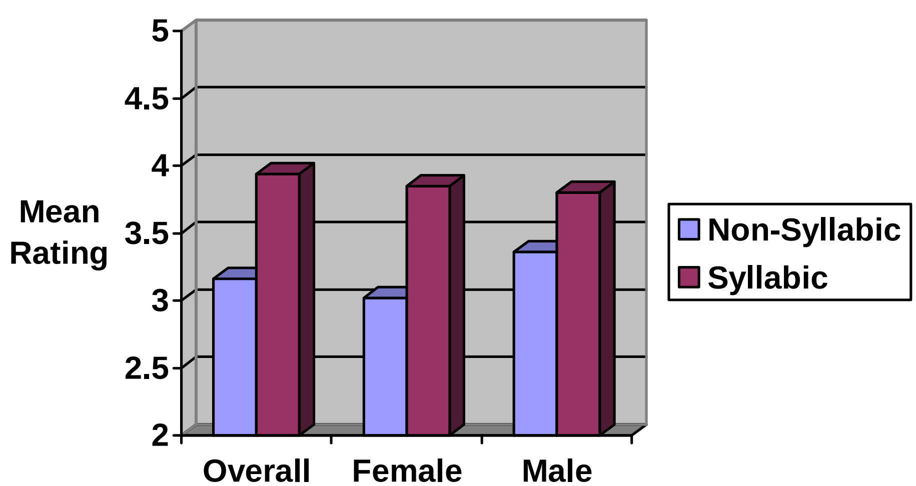
### *5.2.1 Bai Ratings*

Bai listeners made a clear and statistically significant distinction between non-syllabic and syllabic utterances ($F(1, 238) = 207.204$, $p > .001$, $r = .327$), as reflected in overall mean ratings of 3.16 for non-syllabic and 3.94 for syllabic utterances, respectively. As shown in Table 5-6 and Figure 5-6 below, mean ratings of female and male participants differed slightly, in that female ratings were slightly lower than male ratings for both utterance categories. However, the gap between mean ratings for non-syllabic and syllabic utterances was approximately the same for both genders, at 0.83 for women, and 0.72 for men, and gender differences in ratings were not significant ($F$(1, 238) = .504, $p$ > .5, $r$ = .001). Overall, these findings suggest that for the Bai participants, utterance type was a stronger and more consistent factor in determining ratings of infant vocalizations than laryngeal quality.

**Table 5-6. Bai ratings by utterance type, overall and by gender.**

| Utterance Type | Overall Mean (s.d.) (n=40) | Female Mean (s.d.) (n=24) | Male Mean (s.d.) (n=16) |
|---|---|---|---|
| Non-Syllabic | 3.16 (1.47) | 3.02 (1.43) | 3.36 (1.50) |
| Syllabic | 3.94 (1.35) | 3.85 (1.33) | 4.08 (1.35) |

**Figure 5-6. Bai mean ratings of utterance type, overall and by gender.**



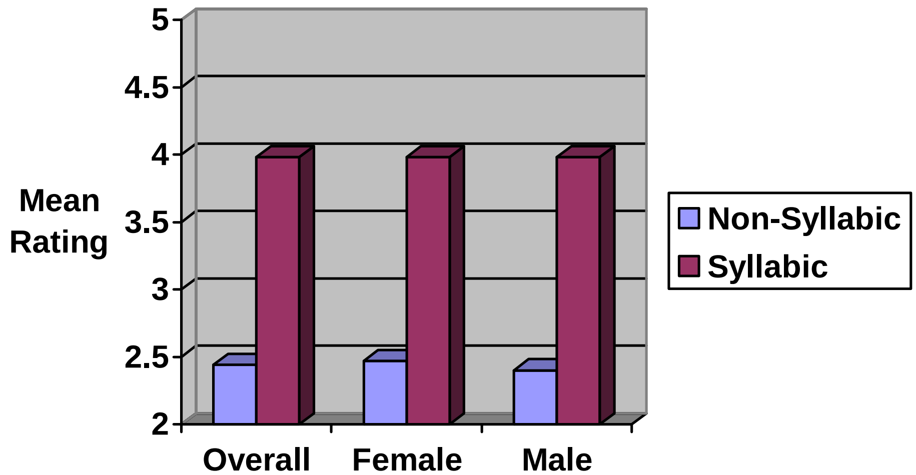
### *5.2.2 English Ratings*

As shown in Table 5-7 and Figure 5-7, English listeners' mean ratings reflect a highly significant preference for syllabic infant vocalizations ($F(1, 238) = 866.605$, $p >$ .001, $r = .749$), as reflected in a mean overall rating of 3.98, versus a mean overall rating of 2.44 for non-syllabic vocalizations. As seen in the ratings for laryngeal categories, mean female and male ratings were highly similar, showing no significant differences ($F(1, 238) = 2.255$, $p > .5$, $r = .008$). The mean female rating for non-syllabic utterances was 2.47, while the mean male rating for the same utterance type was 2.40. Mean female and male ratings for syllabic utterances were identical, at 3.98. Moreover, the range of ratings was similar between the two genders, as reflected in female and male standard deviations of 1.19 and 1.11 for non-syllabic utterances, respectively. The corresponding figures for syllabic vocalizations are also similar, at 1.02 and 0.97. The mean ratings by utterance type, together with the mean ratings by laryngeal quality, provide further evidence that the English listeners, whether male or female, tended to evaluate the importance of the various infant vocalizations included in the listening task in similar ways.

**Table 5-7. English ratings of utterance type, overall and by gender.**

| Utterance Type | Overall Mean (s.d.) (n=40) | Female Mean (s.d.) (n=19) | Male Mean (s.d.) (n=21) |
|---|---|---|---|
| Non-Syllabic | 2.44 (1.15) | 2.47 (1.19) | 2.40 (1.11) |
| Syllabic | 3.98 (0.99) | 3.98 (1.02) | 3.98 (0.97) |

**Figure 5-7. English mean ratings of utterance type, overall and by gender.**




### *5.2.3 Cross-Linguistic Comparison*

The ratings for Bai and English listeners show that both groups of participants made a clear distinction between non-syllabic and syllabic vocalizations, although, as reflected in the effect sizes reported above ($r$ = .749 for English listeners, versus .327 for Bai listeners), the English listeners made a sharper distinction between the two utterance types than did the Bai listeners. While the mean rating for syllabic utterances is similar for both groups, at 3.94 and 3.98 respectively, the Bai listeners tend to rate non-syllabic utterances more highly than do English listeners. As show in Table 5-8 and Figure 5-8 below, the mean English rating for non-syllabic utterances was 2.44, 1.54 points less than the mean syllabic rating. By contrast, the mean Bai rating for non-syllabic utterances was 3.16, only 0.78 points less than the non-syllabic rating. This difference suggests that while both groups place considerable weight on utterance type in judging the relative significance of different infant vocalizations in language development, English listeners weight this factor more heavily than Bai listeners, across genders. (See Tables 5-9 and 5-10 for a comparison of gender across languages.)

**Table 5-8. Bai and English ratings of utterance type, overall.**

| Utterance Type | Bai Mean (s.d.) (n=40) | English Mean (s.d.) (n=40) |
|---|---|---|
| Non-Syllabic | 3.16 (1.47) | 2.44 (1.15) |
| Syllabic | 3.94 (1.35) | 3.98 (0.99) |

**Figure 5-8. Mean ratings of utterance type, by language.**

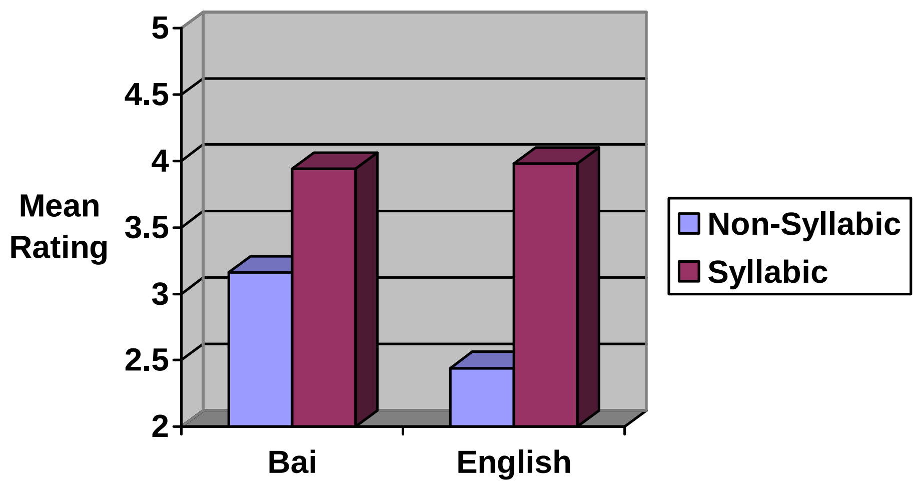



**Table 5-9. Female ratings of utterance type, by language.**

| Utterance Type | Bai Female Mean (s.d.) (n=24) | English Female Mean (s.d.) (n=19) |
|---|---|---|
| Non-Syllabic | 3.02 (1.43) | 2.47 (1.19) |
| Syllabic | 3.85 (1.33) | 3.98 (1.02) |

**Table 5-10. Male ratings of utterance type, by language.**

| Utterance Type | Bai Male Mean (s.d.) (n=16) | English Male Mean (s.d.) (n=21) |
|---|---|---|
| Non-Syllabic | 3.36 (1.50) | 2.40 (1.11) |
| Syllabic | 4.08 (1.35) | 3.98 (0.97) |

In evaluating the significance of the different weight placed on syllabic versus non-syllabic utterances between languages, it is important to consider whether Bai listeners simply rated all the utterances more highly than the English listeners, and if so, whether this trend reflects a failure to complete the task as directed, or an intrinsic property of the Bai population's perception of the relative importance of non-syllabic and syllabic vocalizations in the process of language development. Bai listeners' mean rating across items is 3.55 (s.d. 1.46), somewhat higher than English listeners' mean rating for all utterances, at 3.21 (s.d. 1.32). Listeners from both language groups were asked to use all values on the five-point scale. Eight Bai participants (four male, four female) and six English participants (three male, three female) failed to use the full range of the scale, but in all cases, these individuals did not select '1' for any item. Thus, the upward bias is relatively evenly distributed across the two language groups, and the cross-linguistic difference in the ratings of non-syllabic and syllabic utterances likely reflects a linguistic and/or cultural difference between the Bai and English participants. This issue will be discussed in more detail when reviewing the results of the qualitative component of the listening task in Section 5.4 and in Chapter 6, where the relationship between adult perceptions of infant vocalizations and patterns in infants' production will be considered at length.

### *5.2.4 Summary*

In this section, the mean ratings by utterance type were presented for Bai and English listeners of both genders. These findings show that women and men from both language groups make a clear distinction between non-syllabic and syllabic utterances when judging the relative importance of these utterances to language development. However, while participants from both language groups give higher ratings to syllabic vocalizations than to non-syllabic vocalizations, English listeners accord less significance to non-syllabic utterances than do Bai listeners, a pattern which may reflect the specific demands of language development in Bai versus English.

## 5.3 Laryngeal Voice Quality in Non-Syllabic Utterances

In Section 5.2, ratings for non-syllabic utterances by laryngeal quality will be reviewed. As in previous sections of this chapter, the analysis will focus on identifying universal and language-specific patterns in ratings of the two language groups, overall and by gender, where relevant. Section 5.2.1 provides an overview of the mean ratings for sub-categories of non-syllabic vocalizations, by language. Sections 5.2.2, 5.2.3, and 5.2.4 provide a detailed breakdown of the ratings for individual items within the sub-categories constricted non-syllabic, dynamic non-syllabic, and unconstricted non-syllabic utterances, by language and gender. The latter analysis is intended to account for some of the factors that generate the considerable variability in the sub-category ratings for listeners of both language groups.

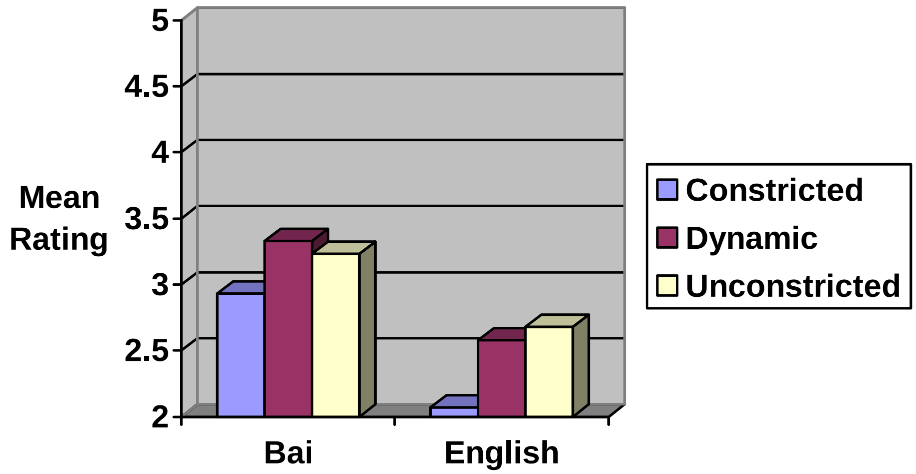
### *5.3.1 Sub-Category Comparisons*

Table 5-11 and Figure 5-9 below show the mean ratings for the sub-categories constricted non-syllabic, dynamic non-syllabic, and unconstricted non-syllabic, by language. A one-way ANOVA showed significant differences between English and Bai listeners' ratings of constricted non-syllabic ($F(1, 478) = 89.269$, $p > .001$), dynamic non-syllabic (*F*(1, 478) = 66.008, *p* > .001), and unconstricted non-syllabic (*F*(1, 478) = 7.252, *p* > .001) utterances. Mean Bai ratings for constricted, dynamic, and unconstricted utterances are 2.93, 3.33, and 3.23, respectively, compared to 2.07, 2.58, and 2.68 for the English listeners. However, the gap between constricted non-syllabic and the other two non-syllabic sub-categories is more marked for the English listeners than for the Bai listeners, at 0.51 versus 0.3, respectively. Thus, while listeners from both language groups accord the least significance to constricted non-syllabic utterances in the process of language development, the English listeners would appear to grant it less importance than the Bai listeners. As previously discussed, overall, Bai listeners' ratings for non-syllabic vocalizations are higher, on average, than English listeners' ratings of non-syllabic utterances. Also, the variability within the ratings is significantly greater among the Bai listeners for every non-syllabic sub-category than that seen among the English listeners, as reflected in the standard deviations reported in Table 5-11. Among the English listeners, there is least variability in ratings of constricted non-syllabic utterances, while for the Bai listeners, this sub-category produces the greatest variability in ratings. In other words, English listeners are relatively convinced of the unimportance of constricted non-syllabic utterances in language development, while Bai listeners disagree widely in their evaluation of the significance of this type of utterance. This difference will be explored more fully in the discussion of the qualitative component of the listening task.

**Table 5-11. Ratings of non-syllabic utterances, by laryngeal quality and language.**

| Laryngeal Quality | Bai Mean Rating (s.d.) | English Mean Rating (s.d.) |
|---|---|---|
| Constricted | 2.93 (1.518) | 2.07 (1.067) |
| Dynamic | 3.33 (1.415) | 2.58 (1.179) |
| Unconstricted | 3.23 (1.440) | 2.68 (1.221) |

**Figure 5-9. Ratings of non-syllabic utterances, by laryngeal quality and language.**




### *5.3.2 Constricted Non-Syllabic*

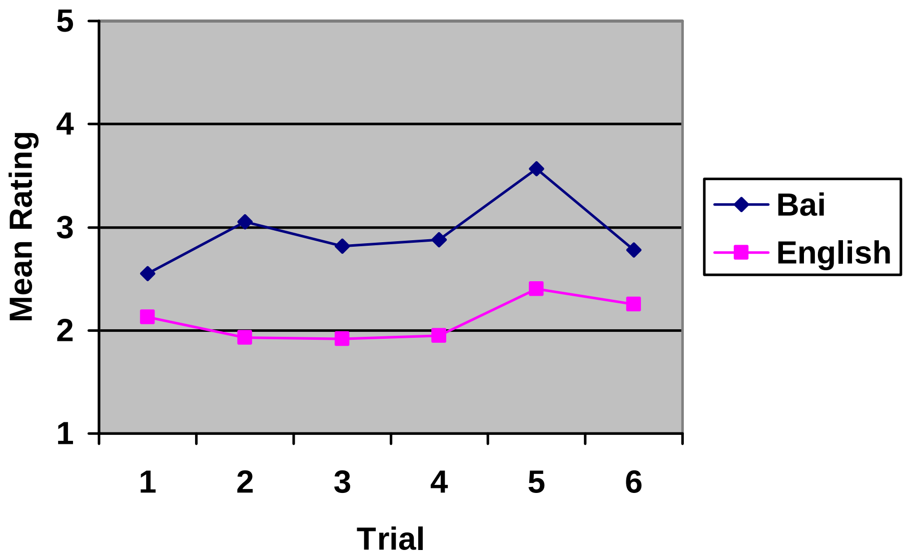
*5.3.2.1 Cross-Linguistic Comparison*

Table 5-12 below shows the mean ratings, by trial, for the constricted non-syllabic utterances. Given the fact that Bai listeners' mean ratings for non-syllabic vocalizations are generally higher than English listeners' mean ratings for these utterances, it is more useful to compare the relative ranking of individual items in a visual medium that shows common and divergent upward and downward trends between the two languages. These trends are plotted in Figure 5-10 below.

**Table 5-12. Ratings of constricted non-syllabic utterances, by trial and language.**

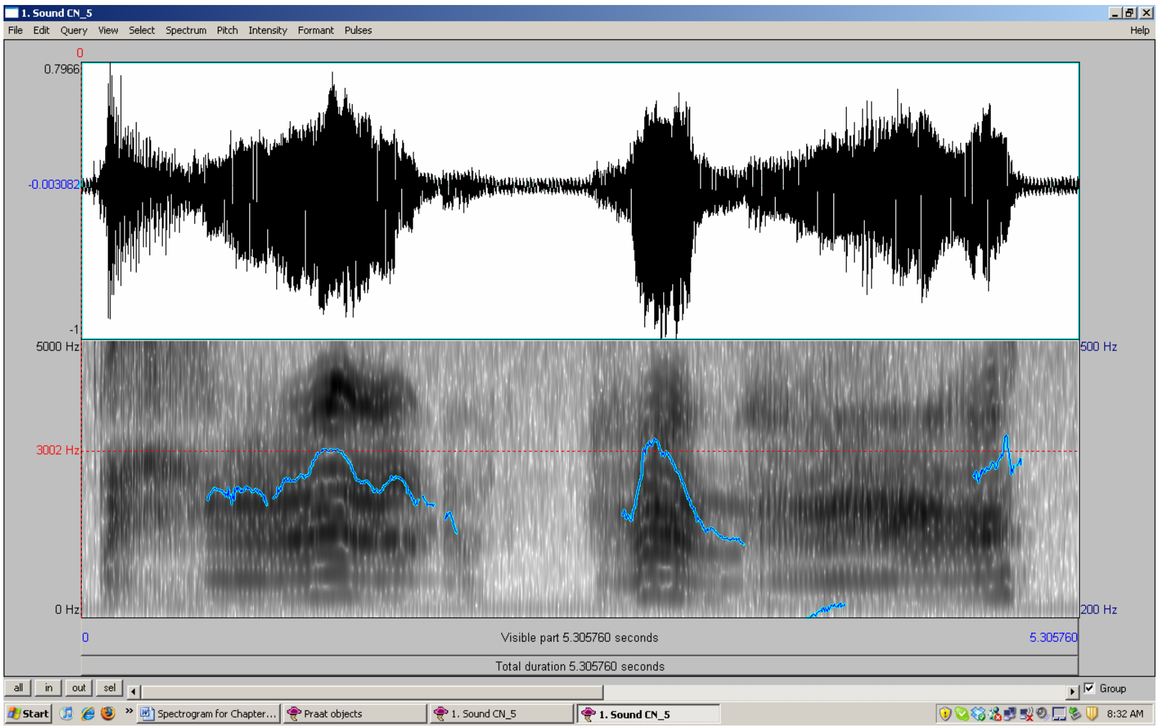
| Trial (Infant's Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| CN-1 (Bai) | 2.55 (1.358) | 2.13 (1.114) |
| CN-2 (Bai) | 3.05 (1.535) | 1.93 (0.888) |
| CN-3 (Bai) | 2.82 (1.534) | 1.92 (0.944) |
| CN-4 (English) | 2.88 (1.539) | 1.95 (0.959) |
| CN-5 (English) | 3.57 (1.583) | 2.40 (1.317) |
| CN-6 (English) | 2.78 (1.476) | 2.25 (0.981) |

**Figure 5-10. Ratings of constricted non-syllabic utterances, by trial and language.**



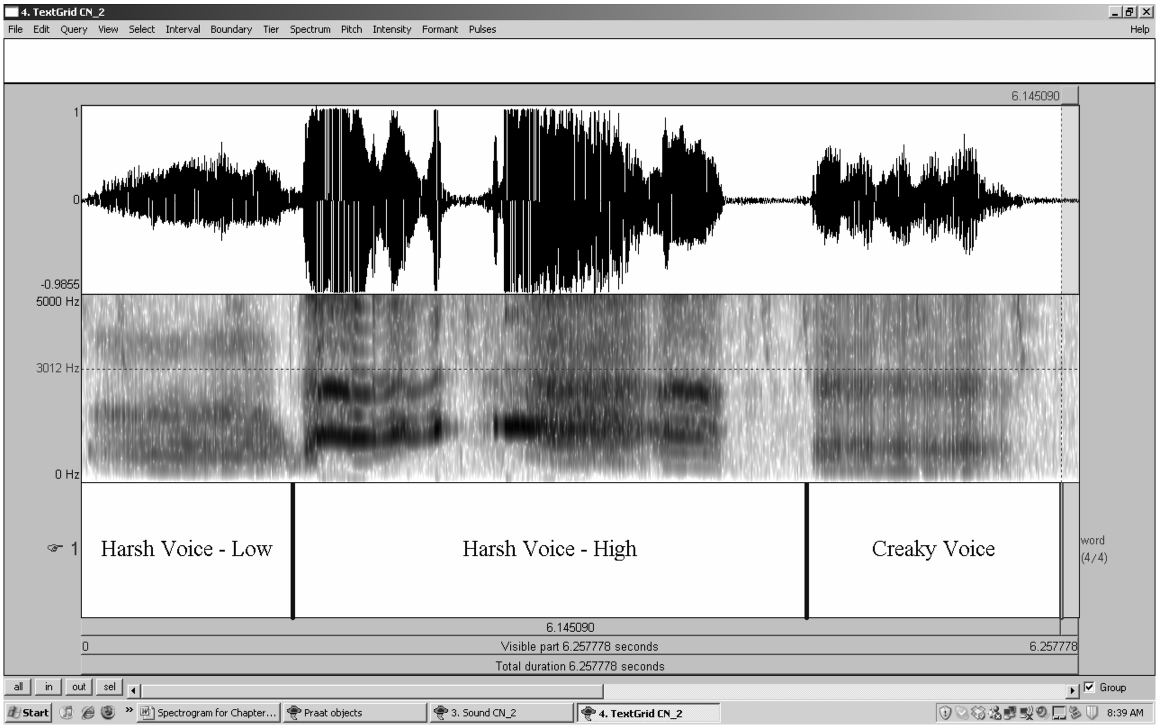

As illustrated in Figure 5-10 above, Bai and English listeners agree in their relative judgments of the importance of trials 1, 3, 4, 5, and 6, but contrast in their relative evaluations of trial 2. Both groups of listeners assign trial 5 the highest mean rating among the constricted non-syllabic vocalizations. Like most of the other utterances in this sub-category, trial 5 is produced with low-pitched harsh voice. However, unlike some of the other constricted non-syllabic vocalizations, this utterance features the production of some smooth pitch contours within this harsh setting, as shown in the spectrogram in Figure 5-11. Listeners from both language groups appear to respond favourably to this feature.

**Figure 5-11. Spectrogram of constricted non-syllabic, trial 5.**




By contrast, the two language groups differ in their relative evaluations of trial 2. For English listeners, this sound is evaluated as among the least important of the constricted non-syllabic utterances for language development, while for Bai listeners, this sound receives the second-highest rating among the utterances in this sub-category. As shown in Figure 5-12, this utterance is produced with very harsh voice. In this utterance, the infant begins in low-pitched harsh voice in a level pitch, and then makes an abrupt transition to high-pitched harsh voice, returning to low-pitched creaky voice at the end of the vocalization. Based on their mean ratings of this utterance, Bai listeners appear to favour the ability to make rapid transitions in pitch and voice quality, while English listeners have a relative aversion to such sounds.

**Figure 5-12. Spectrogram of constricted non-syllabic, trial 2.**




*5.3.2.2 Gender Differences*

There are noticeable, though non-significant ($F(1, 238) = 5.256$, $p > .5$), gender differences in the perceptions of the constricted non-syllabic sounds among the Bai listeners, including for trial 2, discussed above. As shown in Table 5-13, Bai women and men are relatively congruent in their evaluations of trials 1, 3, and 5, but differ significantly in their judgments of trials 2, 4, and 6, with Bai men assigning higher ratings to these sounds than Bai women. Trials 2, 4, and 6 are similar in that they are examples of infants who produce rapid shifts in pitch or voice quality. For example, in trial 6, illustrated in Figure 5-14, the infant produces a long stretch of whispery voice, followed by an abrupt transition to high-pitched harsh voice.

**Table 5-13. Bai ratings of constricted non-syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=24) | Male Rating (s.d.) (n=16) |
|---|---|---|
| CN-1 (Bai) | 2.68 (1.376) | 2.33 (1.345) |
| CN-2 (Bai) | 2.80 (1.607) | 3.47 (1.356) |
| CN-3 (Bai) | 2.92 (1.613) | 2.69 (1.448) |
| CN-4 (English) | 2.58 (1.472) | 3.31 (1.580) |
| CN-5 (English) | 3.58 (1.501) | 3.56 (1.750) |
| CN-6 (English) | 2.33 (1.404) | 3.44 (1.365) |

**Figure 5-13. Bai ratings of constricted non-syllabic utterances, by trial and gender.**

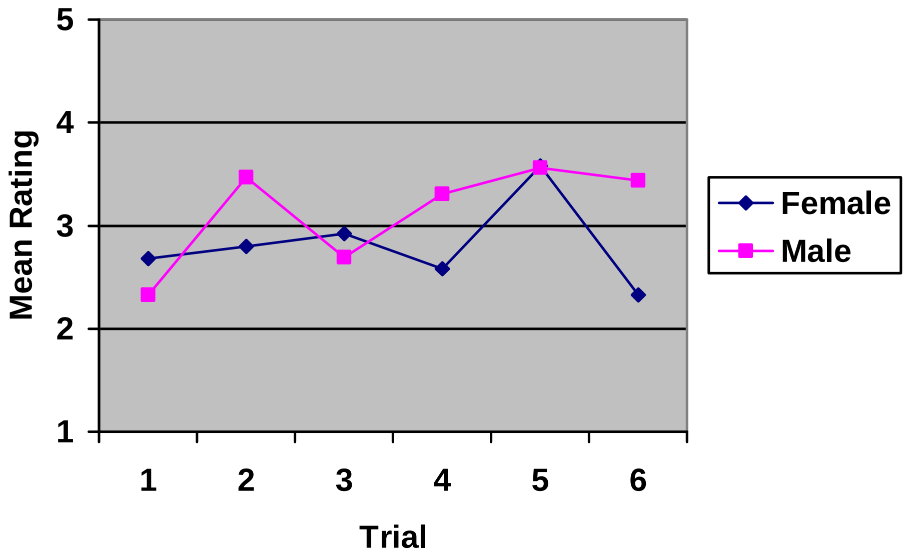
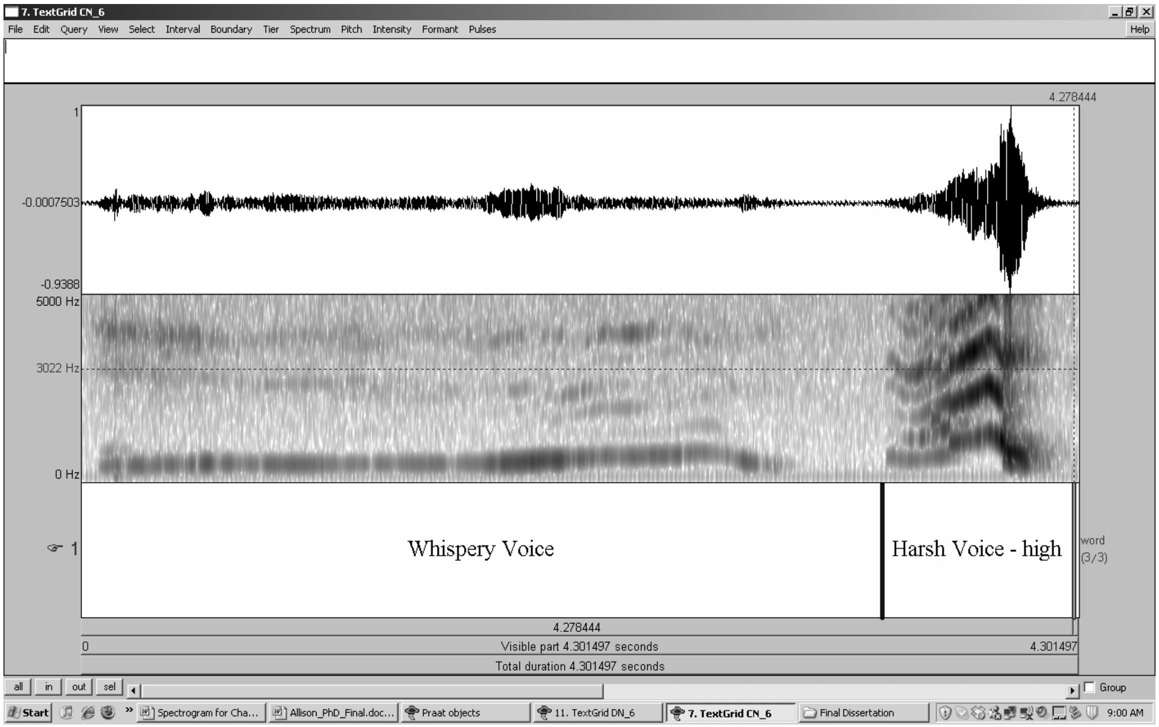


**Figure 5-14. Spectrogram of constricted non-syllabic, trial 6.**

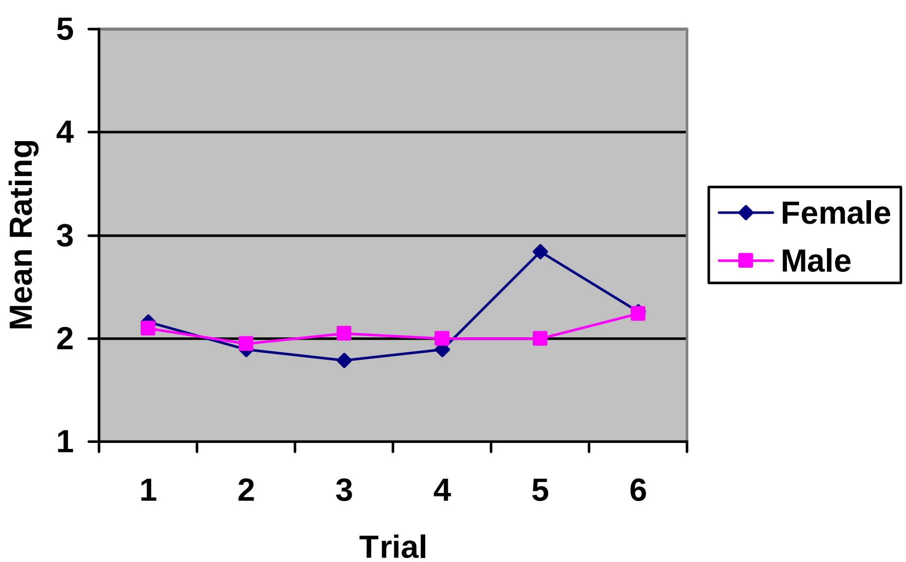



Gender differences among the English listeners are not significant ($F(1, 238) = 1.696$, $p > .5$), though English women and men differ noticeably in their ratings for trial 5, illustrated in Figure 5-11 above, the constricted utterance that featured smooth pitch contours. As shown in Table 5-14 below and Figure 5-15, women's mean ratings for this sound are significantly higher than men's, at 2.84 versus 2.00. While the standard deviation for ratings of this sound is wider than for the other constricted non-syllabic sounds, it is nonetheless of interest that among the constricted non-syllabic sounds, some English women should prefer this utterance. On the whole, however, despite the variability in pitch and voice quality in the constricted non-syllabic utterances, most English women and men do not make many distinctions among them, judging from their mean ratings. The combination of marked laryngeal constriction and non-syllabicity seems to make English listeners most averse to these utterances among the sounds included in this listening task. Most English listeners consistently judge them to be of low importance in learning to speak the target language.

**Table 5-14. English ratings of constricted non-syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| CN-1 (Bai) | 2.16 (1.015) | 2.10 (1.221) |
| CN-2 (Bai) | 1.89 (0.809) | 1.95 (0.973) |
| CN-3 (Bai) | 1.79 (0.855) | 2.05 (1.024) |
| CN-4 (English) | 1.89 (0.809) | 2.00 (1.095) |
| CN-5 (English) | 2.84 (1.425) | 2.00 (1.095) |
| CN-6 (English) | 2.26 (1.098) | 2.24 (0.889) |

**Figure 5-15. English ratings of constricted non-syllabic, by trial and gender.**



### *5.3.3 Dynamic Non-Syllabic*

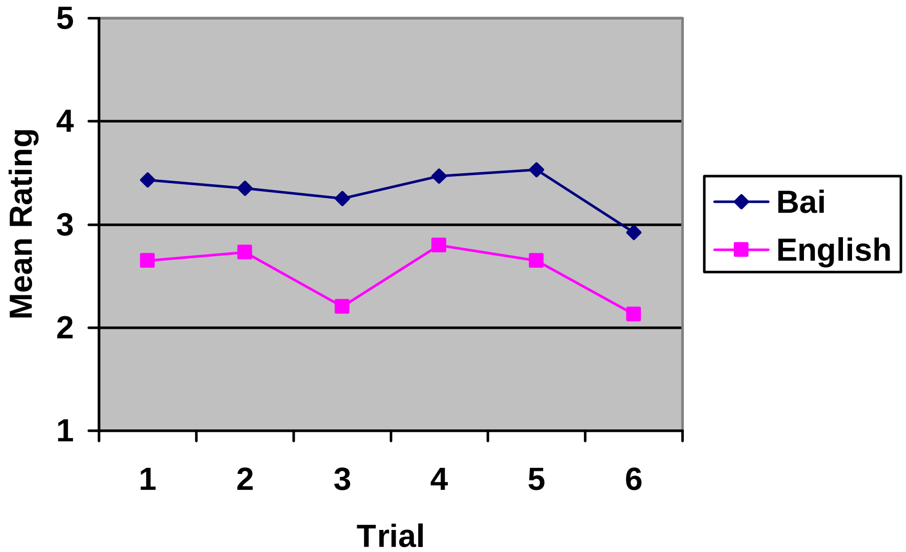
#### *5.3.3.1 Cross-Linguistic Comparison*

Table 5-15 lists the mean ratings of dynamic non-syllabic utterances, by trial and language. Bai listeners tend to assign higher ratings to these utterances than do English listeners. For the Bai listeners, mean ratings of dynamic non-syllabic utterances fall between ‘3’ and ‘4’ for all but one of these vocalizations, while for the English listeners, the mean rating for all six of these sounds falls between ‘2’ and ‘3.’ However, as illustrated in Figure 5-16 below, the relative ranking of the six trial items is approximately the same for both language groups. Moreover, both groups assign higher average ratings to dynamic non-syllabic utterances than to constricted non-syllabic utterances, though the distinction between the two sub-categories is more marked for the English listeners than for the Bai listeners.

**Table 5-15. Ratings of dynamic non-syllabic utterances, by trial and language.**

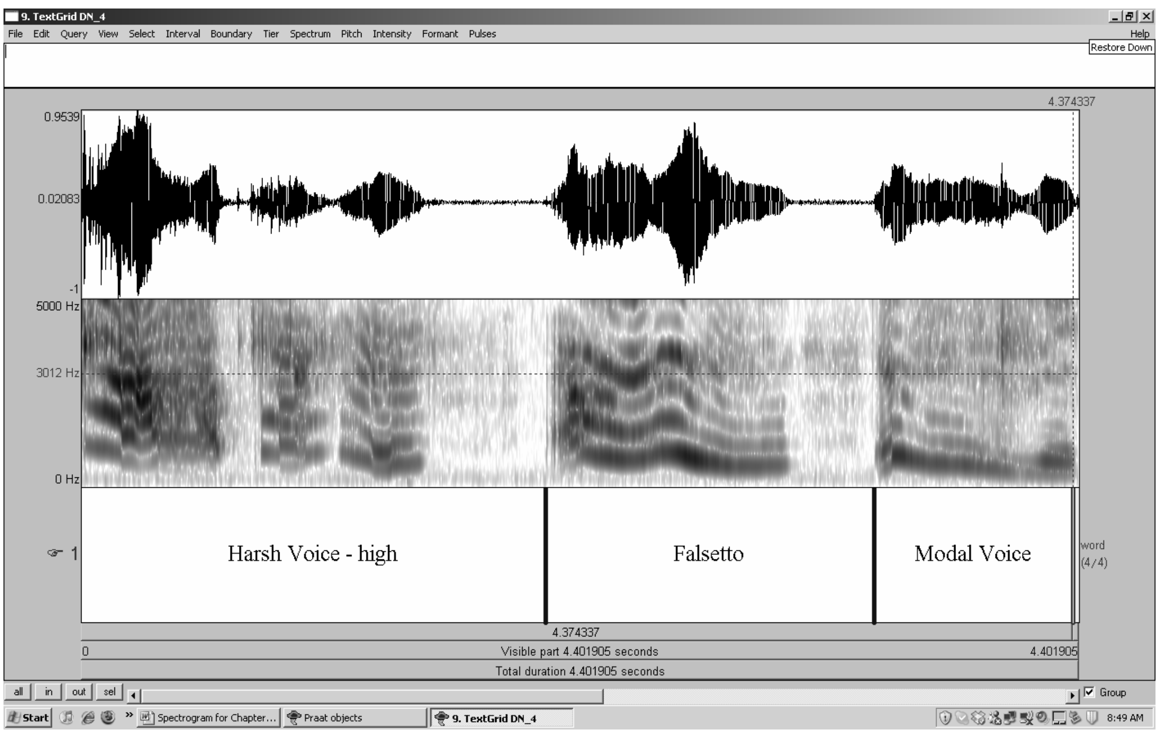
| Trial (Infant’s Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| DN-1 (Bai) | 3.43 (1.430) | 2.65 (1.145) |
| DN-2 (Bai) | 3.35 (1.406) | 2.73 (1.062) |
| DN-3 (Bai) | 3.25 (1.481) | 2.20 (1.114) |
| DN-4 (English) | 3.47 (1.281) | 2.80 (1.305) |
| DN-5 (English) | 3.53 (1.467) | 2.65 (1.189) |
| DN-6 (English) | 2.92 (1.421) | 2.13 (0.992) |

**Figure 5-16. Ratings of dynamic non-syllabic utterances, by trial and language.**




The dynamic utterances included in this sub-category are all examples of what would be referred to in the literature as “vocal play,” (Oller, 2000, 2004), which in the context of this study, refers to a vocalization in which an infant produces a rapid alternation between a constricted and unconstricted voice quality setting, often in concert with wide swings in pitch. Judging from auditory impressions and visual inspection of spectrograms, generally speaking, the constricted portion of dynamic non-syllabic utterances is not usually as constricted as the degree of constriction featured in many constricted non-syllabic utterances, including the constricted utterances included in the group of constricted non-syllabic vocalizations that were used in this listening task. All the utterances included in the current sample involved the infant moving from a low- or mid-pitched setting into a high-pitched setting, or vice versa. Usually, the high-pitched portion of the utterance was constricted, and the mid- or low-pitched portion was unconstricted, but sometimes the high-pitched portion included unconstricted, i.e., falsetto, quality, and the low-pitched portion involved varying degrees of harsh voice, the least constricted of these degrees featuring larynx raising but no aperiodicity in vocal fold vibration or aryepiglottic trilling. Figure 5-17 is a representative example of the dynamic non-syllabic utterances included in this study. The utterance begins in high-pitched harsh voice. Towards the middle of the sound, the infant moves into falsetto. Near the end of the utterance, pitch lowers, and the infant produces modal voice. While there are variations in the ratings across items, these variations appear to stem primarily from gender differences, to be discussed in Section 5.2.3.2. below.

**Figure 5-17. Spectrogram of dynamic non-syllabic, trial 4.**



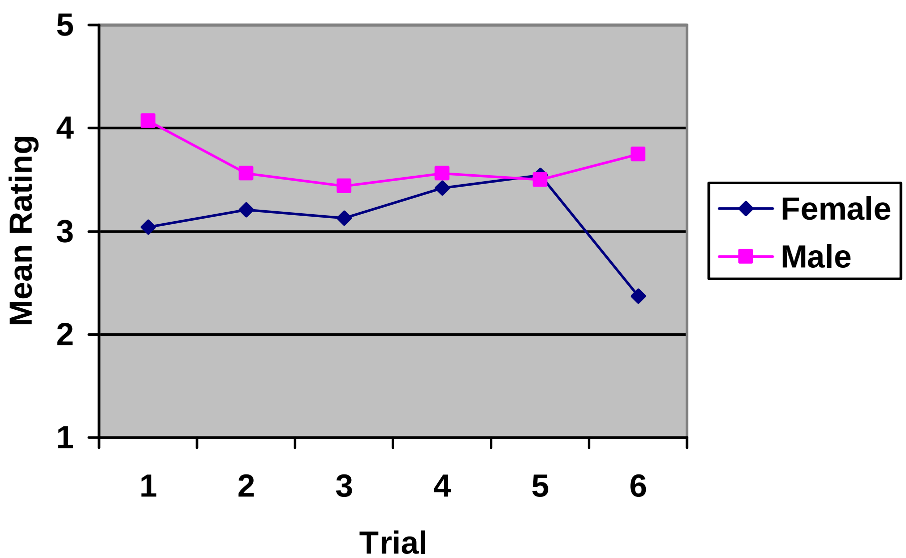
*5.3.3.2 Gender Differences*

As shown in Table 5-16 and Figure 5-18 below, Bai men tend to assign higher ratings to dynamic non-syllabic utterances than do Bai females, a difference which is statistically significant ($F(1, 238) = 14.40$, $p > .05$). Given that among the constricted non-syllabic sounds, Bai males gave higher ratings to utterances that exhibited a rapid change in pitch or degree of constriction than did Bai females, the Bai men's higher ratings of dynamic non-syllabic utterance may reflect the influence of this same factor, given that rapid alternations in pitch and quality are the hallmark of this type of utterance.

**Table 5-16. Bai ratings of dynamic non-syllabic utterances, by trial and gender.**

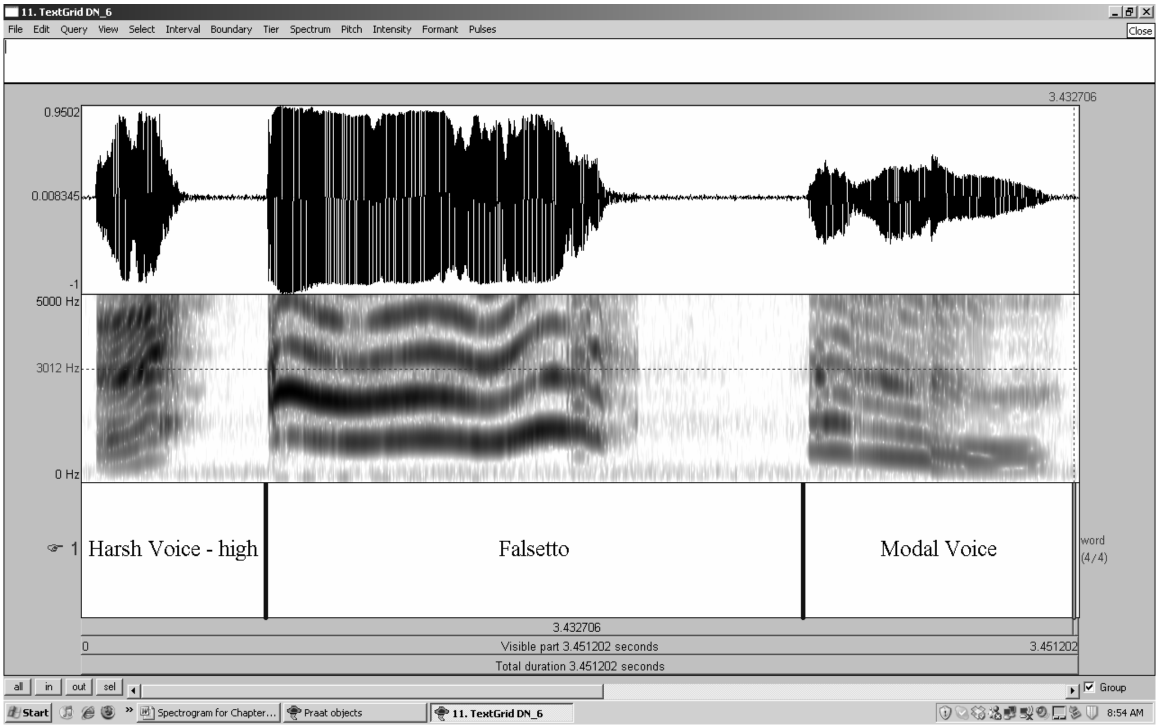
| Trial (Infant's Language) | Female Rating (s.d.)<br>(n=24) | Male Rating (s.d.)<br>(n=16) |
|---|---|---|
| DN-1 (Bai) | 3.04 (1.306) | 4.07 (1.438) |
| DN-2 (Bai) | 3.21 (1.318) | 3.56 (1.548) |
| DN-3 (Bai) | 3.13 (1.424) | 3.44 (1.590) |
| DN-4 (English) | 3.42 (1.248) | 3.56 (1.365) |
| DN-5 (English) | 3.54 (1.250) | 3.50 (1.789) |
| DN-6 (English) | 2.37 (1.245) | 3.75 (1.291) |

**Figure 5-18. Bai ratings of dynamic non-syllabic utterances, by trial and gender.**



The relatively large gap between the mean ratings of women and men on trials 1 and 6 is harder to explain, as most sounds included in this sub-category are relatively similar in terms of pitch range and voice quality. However, the samples in this sub-category do vary in length. Trials 1 and 6 are the shortest of the six sounds, at 3.72 sec and 3.45 sec, respectively. By comparison, trials 2 and 4 are 4.64 sec and 4.40 sec, respectively, and trials 3 and 5, the longest in the sample, are 5.44 sec and 5.55 sec, respectively. Judging from the gender ratings of all six utterances, length systematically affects the ratings of Bai women and men. When a rapid alternation in pitch or voice quality occurs in a short time frame, male Bai listeners award the sound a higher mean rating, and female Bai listeners assign the sound a lower rating. When an alternation in pitch or voice quality is more gradual, male Bai listeners assign it a lower rating, and female Bai listeners give it a higher rating. Figure 5-19 below, a spectrogram of trial 6, illustrates the kind of transition to which the male Bai listeners responded favourably, and to which the female Bai listeners assigned lower ratings. The first part of the utterance is in harsh high voice, following which there is an abrupt shift to falsetto, shown in the middle part of the utterance. At the end of the utterance, the infant shifts into mid and low pitch in modal voice. Possible reasons for such differences in the perceptions of Bai women versus Bai men will be explored in the qualitative component of the perceptual study.

**Figure 5-19. Spectrogram of dynamic non-syllabic, trial 6.**




English ratings for dynamic non-syllabic utterances do not differ significantly ($F(1, 238) = .017$, $p > .1$), as evidenced in Table 5-17 and Figure 5-18, though the gap between women's and men's ratings are wider on trial 3 than for other items in this sub-category, at 1.95 for women, versus 2.43 for men. As shown in Figure 5-21 below, a wide-band spectrogram of trial 3, this utterance starts out in very high-pitched harsh voice that continues into the second portion of the utterance, though at a somewhat lower pitch. In the third portion of the utterance, the infant shifts into mid-pitched harsh voice before making a steady transition into low-pitched modal voice. Though all the utterances included in the sub-category included high pitch, this utterance featured the highest pitch and the widest pitch range of the sample. While English listeners generally responded more negatively to these features than Bai listeners, English female listeners may have reacted more negatively to them than English male listeners. In common with Bai females, both male and female English listeners responded negatively to trial 6, pictured in Figure 5-18 above, which featured abrupt transitions in pitch and voice quality.

**Table 5-17. English ratings of dynamic non-syllabic utteranaces, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| DN-1 (Bai) | 2.84 (1.214) | 2.48 (1.078) |
| DN-2 (Bai) | 2.58 (0.961) | 2.86 (1.153) |
| DN-3 (Bai) | 1.95 (1.177) | 2.43 (1.028) |
| DN-4 (English) | 2.79 (1.357) | 2.81 (1.289) |
| DN-5 (English) | 2.47 (1.172) | 2.81 (1.209) |
| DN-6 (English) | 2.26 (1.195) | 2.00 (0.775) |

**Figure 5-20. English ratings of dynamic non-syllabic utterances, by trial and gender.**

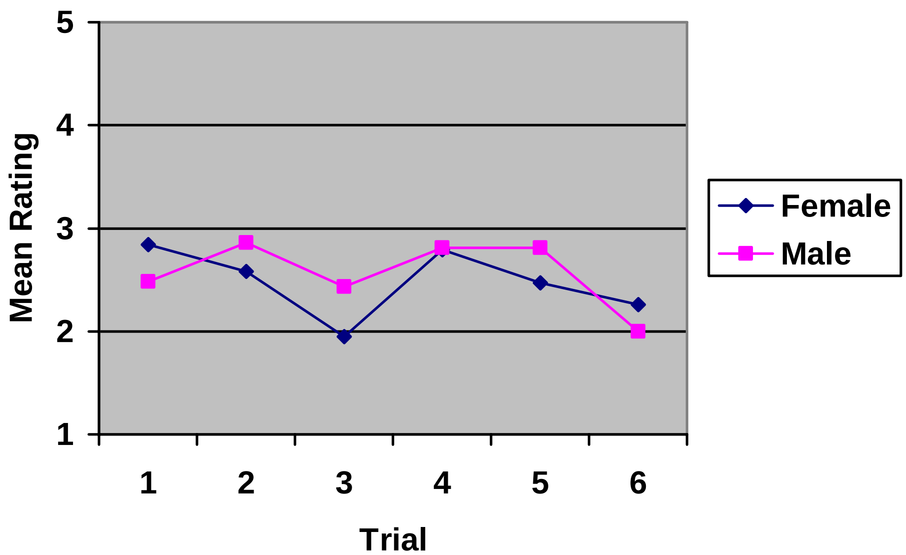
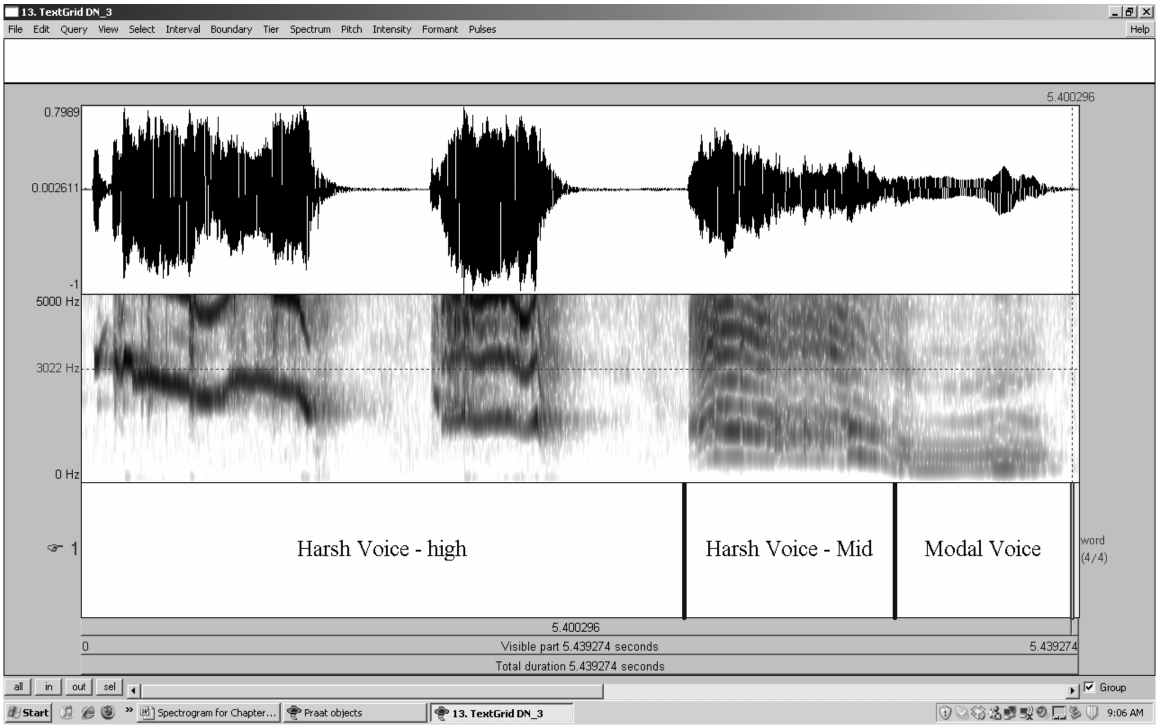


**Figure 5-21. Spectrogram of dynamic non-syllabic, trial 3.**




### 5.3.4 *Unconstricted Non-Syllabic*

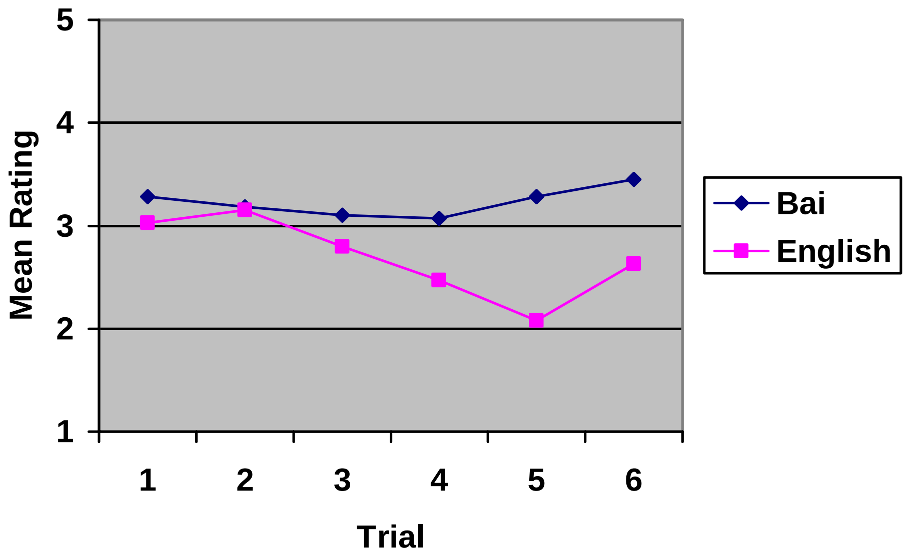
#### *5.3.4.1 Cross-Linguistic Comparison*

The ratings for unconstricted non-syllabic utterances reveal interesting similarities and differences between the two language groups. Table 5-18 and Figure 5-22 show the mean ratings per item by trial. As reflected in these ratings, Bai and English ratings are fairly similar for trials 1, 2, and 3, but differ noticeably on trials 4, 5, and 6. Overall, the Bai ratings for unconstricted non-syllabic utterances are distributed in a similar range to that seen for Bai ratings of dynamic non-syllabic utterances, though the overall mean for dynamic non-syllabic utterances is slightly higher than for unconstricted non-syllabic utterances. By contrast, English ratings for unconstricted non-syllabic sounds fall into two groups. The mean ratings for trials 1, 2, and 3 are at least as high as any of the ratings for the non-syllabic dynamic utterances, with trials 1 and 2 receiving the highest ratings awarded by the English listeners for any of the non-syllabic utterances in any laryngeal category. However, English ratings for trials 4, 5, and 6 fall within the range typically seen for English ratings of non-syllabic utterances. As will be explained below, these differences relate to different uses of pitch and voice quality within unconstricted utterances.

**Table 5-18. Ratings of unconstricted non-syllabic utterances, by trial and language.**

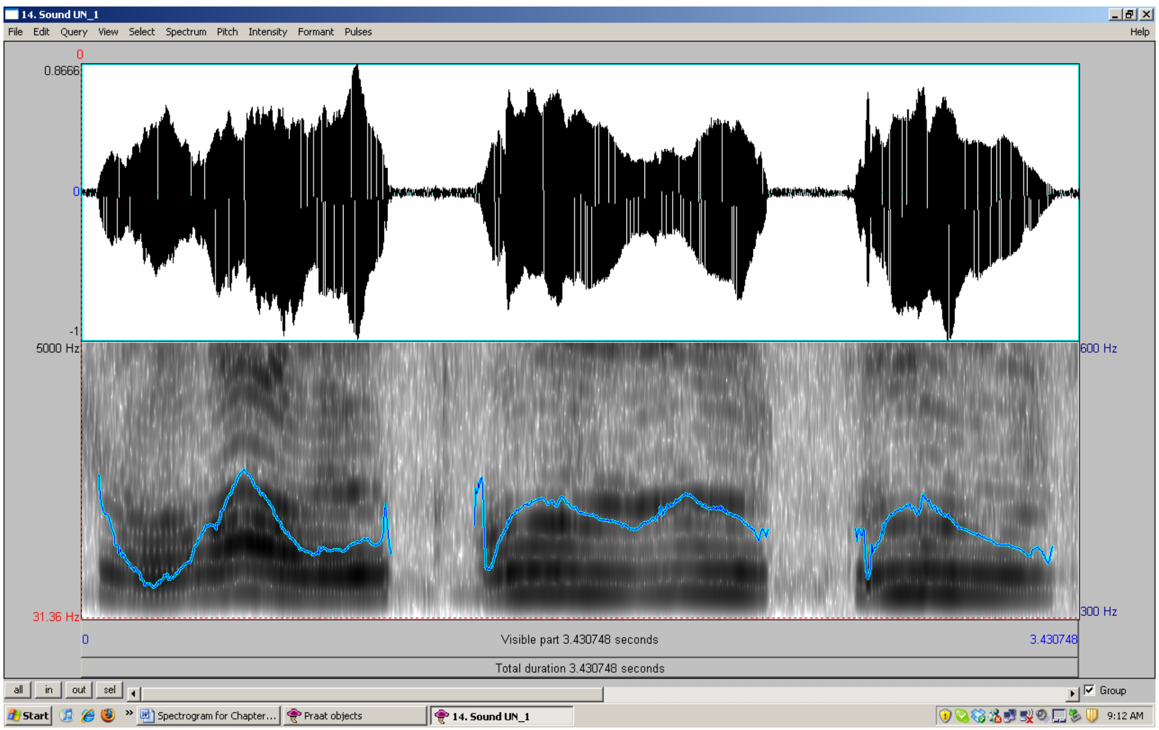
| Trial (Infant's Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| UN-1 (Bai) | 3.28 (1.432) | 3.03 (1.187) |
| UN-2 (Bai) | 3.18 (1.338) | 3.15 (1.001) |
| UN-3 (Bai) | 3.10 (1.598) | 2.80 (1.203) |
| UN-4 (English) | 3.07 (1.542) | 2.47 (1.198) |
| UN-5 (English) | 3.28 (1.467) | 2.08 (0.971) |
| UN-6 (English) | 3.45 (1.300) | 2.63 (1.148) |

**Figure 5-22. English and Bai ratings of unconstricted non-syllabic utterances.**



Trials 1, 2, and 3, to which both language groups assign similar ratings, are all produced in modal voice with smooth pitch contours, within a moderate pitch range. Figure 5-23 below, a wide-band spectrogram of trial 1, illustrates the smooth vocalic quality that occurs in trials 1, 2, and 3.

**Figure 5-23. Spectrogram of unconstricted non-syllabic, trial 1.**




By contrast, trials 4 and 5, while unconstricted, both feature shifts in voice quality and/or pitch. Trial 4 starts out in mid-pitched modal voice, followed by a slightly higher-pitched breathy voice. Trial 5 is a sequence that starts out in falsetto and then shifts into mid-pitched modal voice, spanning a pitch range that is more typically seen in dynamic non-syllabic utterances. Trial 6, while uttered in modal voice throughout, differs from the other sounds in the sample in featuring a series of glottal stops and fricatives that break the sound into syllable-like units that vary in length, and that, therefore, include small episodes of mild laryngeal constriction due to the glottal stops. English listeners award all three sounds lower ratings than Bai listeners do, reserving their lowest ratings for trials 4 and 5, which feature pitch and voice quality variations; the greater the change in pitch, as in trial 5, the lower the rating. By contrast, Bai listeners reserve their highest ratings for trials 5 and 6, appearing to respond favourably to the variations in pitch and quality at the laryngeal level. Thus, for English listeners, unconstricted utterances are favoured when they feature smooth pitch contours, relatively constant voice quality, and a moderate pitch range. While Bai listeners also respond favourably to these features, as a group, they respond more favourably than English listeners to sounds that vary in pitch, pitch range, and voice quality, whether these variations occur within constricted or unconstricted settings.

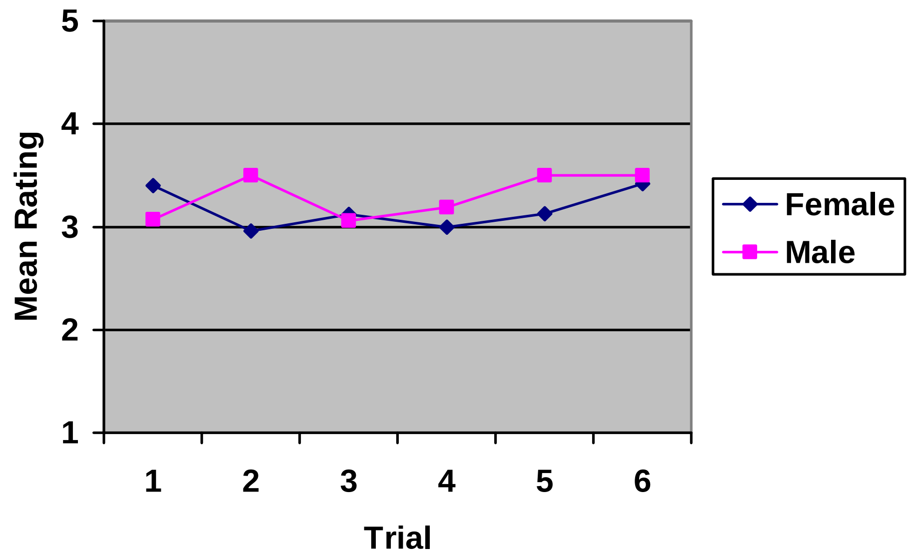
#### *5.3.4.2 Gender Differences*

Based on an analysis of the gender differences in the mean ratings of Bai women and men, shown in Table 5-19 and Figure 5-24 below, it appears that while the aforementioned preference for variations in pitch and voice quality is shared by Bai listeners from both genders, the preference is more marked among Bai male listeners, though gender differences in the ratings of unconstricted non-syllabic utterances are not statistically significant ($F(1, 238) = .951$, $p > .5$).

**Table 5-19. Bai ratings of unconstricted non-syllabic utterances, by trial and gender.**

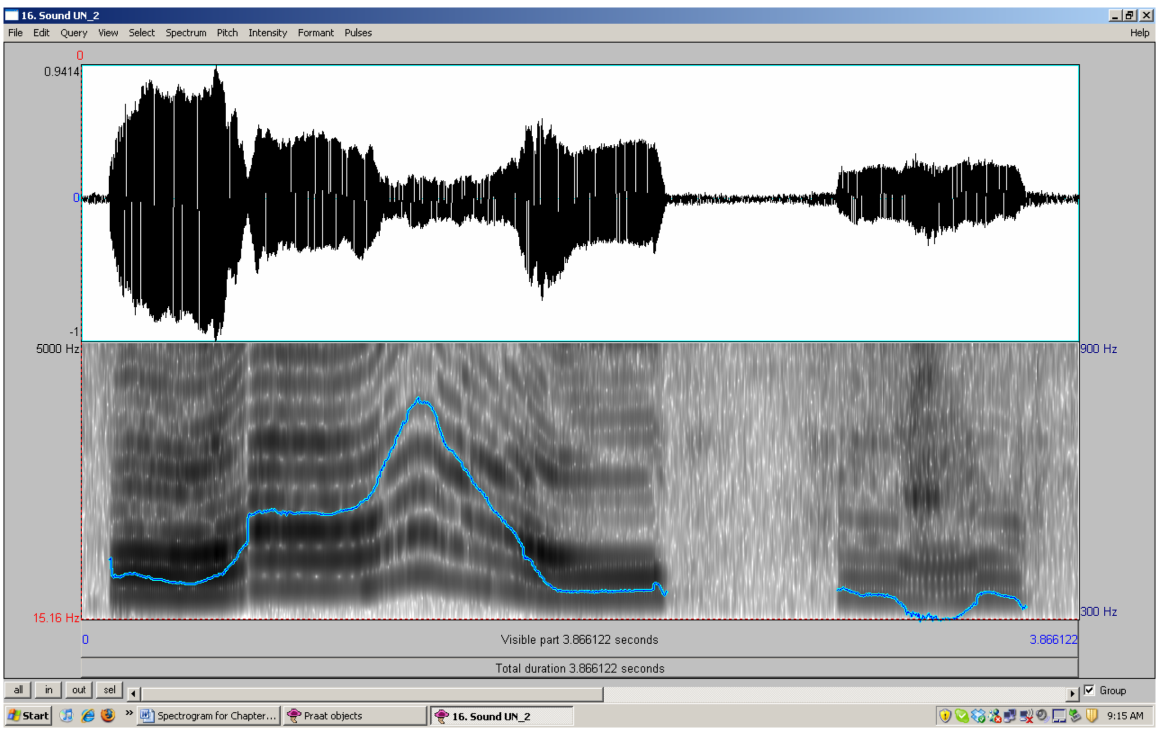
| Trial (Infant's Language) | Female Rating (s.d.) (n=24) | Male Rating (s.d.) (n=16) |
|---|---|---|
| UN-1 (Bai) | 3.40 (1.354) | 3.07 (1.580) |
| UN-2 (Bai) | 2.96 (1.367) | 3.50 (1.265) |
| UN-3 (Bai) | 3.12 (1.650) | 3.06 (1.569) |
| UN-4 (English) | 3.00 (1.504) | 3.19 (1.642) |
| UN-5 (English) | 3.13 (1.424) | 3.50 (1.549) |
| UN-6 (English) | 3.42 (1.349) | 3.50 (1.265) |

**Figure 5-24. Bai ratings of unconstricted non-syllabic utterances, by trial and gender.**




As discussed in Section 5.2.4.1 above, trials 1, 2, and 3 feature modal voicing and smooth pitch contours within a moderate pitch range. However, of these three trials, trial 2 features the widest pitch range. Figure 5-25 below is a wide-band spectrogram of trial 2. As shown in the spectrogram, the first part of this sound includes a smooth rise and fall in pitch; the variation in pitch is more dramatic than that seen in the first part of the utterance depicted in Figure 5-23 above. Bai male listeners award this sound the highest rating of trials 1, 2, and 3. The Bai male preference for wider variations in pitch, as well as voice quality, is also seen in mean ratings for trials 4 and 5. While Bai women and men both rated these sounds higher than English women and men, Bai men rated these sounds higher than Bai women, with the greatest gap between the genders appearing for trial 5, which features the widest pitch range and the most noticeable shift in voice quality (from falsetto to modal) of the utterances included in this sample.

**Figure 5-25. Spectrogram of unconstricted non-syllabic, trial 2.**

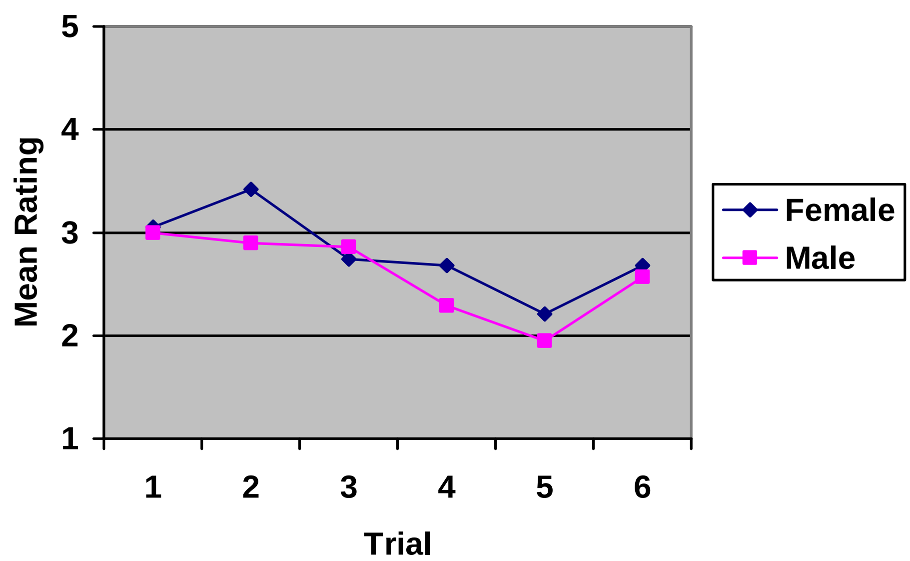



The English ratings for unconstricted non-syllabic utterances also reveal interesting, though statistically non-significant ($F(1, 238) = 1.472$, $p > .1$), gender differences. Table 5-20 and Figure 5-26 show the mean ratings for unconstricted non-syllabic utterances among the English listeners, by gender. While the relative rankings of the six unconstricted non-syllabic sounds is the same for both genders, English females award higher ratings than English males for trials 2, 4, and 5. As highlighted above, among the unconstricted utterances, these sounds are the ones that feature wider variations in pitch and/or voice quality than the other sounds in the sample. Judging by mean ratings, English female listeners show a greater attraction to these sounds than do English male listeners, with the preference being especially marked when the pitch variations occur within a single voice quality. Thus, within the subcategory unconstricted, English female listeners are more similar to Bai listeners than to English male listeners, and in some respects are more similar to Bai males than Bai females.

**Table 5-20. English ratings of unconstricted non-syllabic, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| UN-1 (Bai) | 3.05 (1.129) | 3.00 (1.265) |
| UN-2 (Bai) | 3.42 (0.961) | 2.90 (0.995) |
| UN-3 (Bai) | 2.74 (1.485) | 2.86 (0.910) |
| UN-4 (English) | 2.68 (1.336) | 2.29 (1.056) |
| UN-5 (English) | 2.21 (0.918) | 1.95 (1.024) |
| UN-6 (English) | 2.68 (1.250) | 2.57 (1.076) |

**Figure 5-26. English ratings of unconstricted non-syllabic, by trial and gender.**



### *5.3.5 Summary*

In Section 5.3, I have presented and discussed ratings for non-syllabic utterances, including constricted non-syllabic, dynamic non-syllabic, and unconstricted non-syllabic utterances. The analysis has highlighted a number of patterns that are shared across the two language groups, as well as patterns that differ. The discussion has also included an analysis of gender differences in ratings within language groups. Table 5-21 below summarizes the main points raised in the discussion of non-syllabic ratings between the two language groups.

**Table 5-21. Patterns in ratings of non-syllabic utterances, by language.**

| Sub-Category | Bai Patterns | English Patterns |
|---|---|---|
| Constricted Non-Syllabic | • Lowest-rated among sub-categories, though rated higher than English, and somewhat higher by Bai men than Bai women.<br>• Ratings are positively affected by presence of smooth pitch contours for both genders.<br>• Ratings are positively affected by rapid shifts in pitch and voice quality, particularly for Bai men. | • Lowest-rated among sub-categories.<br>• Ratings are positively affected by presence of smooth pitch contours, particularly for English females.<br>• Ratings are negatively affected by rapid shifts in pitch and voice quality. |
| Dynamic Non-Syllabic | • Most highly rated among non-syllabic utterances.<br>• Ratings of Bai men are positively affected by rapid shifts in pitch and voice quality; ratings of Bai women can be negatively affected by | • Receive considerably higher ratings than constricted non-syllabic.<br>• Ratings are negatively affected by abrupt changes in pitch and voice quality. |

| | | |
|---|---|---|
| | the same factor. | • Ratings are negatively affected by extremes in pitch and pitch range, even if they are distributed gradually throughout the utterance, particularly for English women. |
| Unconstricted Non-Syllabic | • Receive slightly higher ratings than constricted non-syllabic.<br>• Ratings are positively affected by smooth pitch contours in a moderate pitch range.<br>• Ratings are positively affected by intermittent laryngeal constriction, such as glottal stops.<br>• Ratings are positively affected by changes in voice quality, particularly for Bai males. | • Highest-rated among non-syllabic utterances.<br>• Ratings are positively affected by smooth pitch contours and a moderate pitch range, especially for English females, and especially if these occur within a single voice quality.<br>• Ratings are negatively affected by changes in voice quality.<br>• Ratings are negatively affected by intermittent laryngeal constriction, such as the presence of glottal stops. |

Overall, Bai listeners rate non-syllabic utterances higher than do English listeners. Among the Bai listeners, Bai men tend to rate non-syllabic utterances higher than Bai women. Constricted non-syllabic utterances show the greatest differences in ratings between languages. English listeners consistently rate these sounds as among the least important in language acquisition, while Bai listeners rate these sounds as being, on average, only somewhat less important in language acquisition than dynamic and unconstricted non-syllabic utterances. Ratings of dynamic non-syllabic utterances show many common features between the two languages. English listeners rate these sounds as more important in language acquisition than constricted non-syllabic sounds, though they appear to respond more negatively to abrupt changes in pitch and quality than do Bai listeners. In the evaluation of unconstricted non-syllabic sounds, ratings from the two language groups begin to occupy a similar range, though the relative rating of unconstricted sounds can differ widely between the languages, based on features such as variability in pitch and/or voice quality.

The most striking gender differences observed in the analysis of non-syllabic utterances occurred within the Bai listeners. Across all three sub-categories of non-syllabic vocalizations, Bai male listeners showed a preference for sounds with the greatest variability, and most rapid transitions in, voice quality and pitch. While Bai female listeners generally favoured these sounds more than English listeners, they did not rate them as highly as Bai males. Among the English listeners, there were fewer differences in the ratings of women and men. Most of the time, the ratings of English females and males showed a similar pattern within sub-categories. However, within the sub-category of unconstricted non-syllabic vocalizations, English women showed a greater preference for sounds exhibiting variations in pitch and voice quality than did English men. This preference was not strongly manifest in English females' rankings of pitch and voice quality variability within the sub-categories of constricted and dynamic non-syllabic vocalizations. Thus, English women favour variability in pitch and voice quality within a narrower range of vocalizations than do Bai listeners.

## 5.4 Laryngeal Voice Quality in Syllabic Utterances

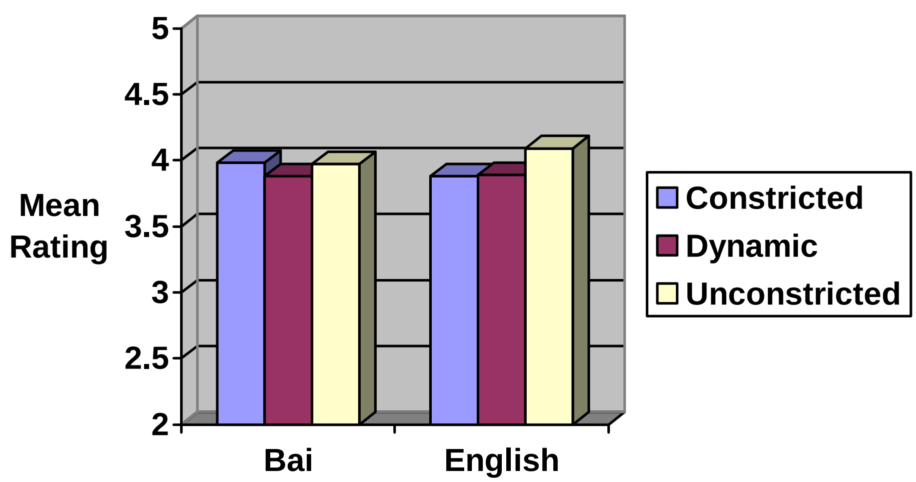
In this section, I review the findings of the perceptual study for syllabic utterances, including the sub-categories constricted syllabic, dynamic syllabic, and unconstricted syllabic. While Bai and English listeners gave significantly different ratings to all sub-categories of non-syllabic utterances, their ratings for syllabic utterances do not differ as significantly. One-way ANOVA analyses showed no significant differences between languages for constricted syllabic utterances ($F(1, 478) = 1.20$, $p > .1$) or dynamic syllabic utterances ($F(1, 478) = 0.19$, $p > .5$), though the two language groups differed significantly in their ratings of unconstricted syllabic utterances ($F(1, 478) = 7.252$, $p > .05$). As shown in the mean sub-category ratings in Table 5-22 and Figure 5-27 below, neither language group shows a strong preference for any laryngeal quality. Mean Bai ratings for constricted, dynamic, and unconstricted syllabic vocalizations are 3.98, 3.88, and 3.97, respectively. The corresponding figures for English listeners are 3.88, 3.89, and 4.09. While English listeners show a slight and statistically significant preference for unconstricted vocalizations, the difference in their mean ratings for this sub-category versus the other two sub-categories of syllabic utterances is not striking. There is, however, considerable variability in the ratings of syllabic vocalizations, particularly among the Bai listeners. To clarify the factors that influence listeners' evaluations of the relative importance of different syllabic utterances to the process of language development in their respective languages (English and Bai), it is necessary to examine individual trials within each sub-category, by language and gender.

Section 5.4.1 presents the findings for constricted syllabic utterances, and Sections 5.4.2 and 5.4.3 outline the findings for dynamic syllabic and unconstricted syllabic vocalizations, respectively. A summary of the findings for syllabic utterances is provided in Section 5.4.4.

**Table 5-22. Ratings of syllabic utterances, by laryngeal quality and language.**

| Laryngeal Quality | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| Constricted | 3.98 (1.330) | 3.88 (1.013) |
| Dynamic | 3.88 (1.371) | 3.89 (1.065) |
| Unconstricted | 3.97 (1.346) | 4.09 (1.162) |

**Figure 5-27. Ratings of syllabic utterances, by laryngeal quality and language.**




### *5.4.1 Constricted Syllabic*

*5.4.1.1 Cross-Linguistic Comparison*

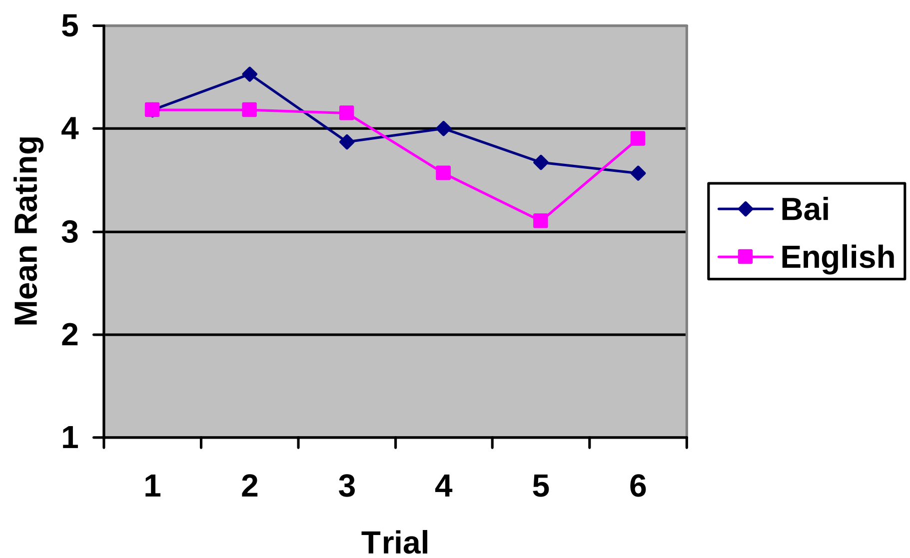
As shown in Table 5-23 and Figure 5-28, Bai and English ratings show a number of common and divergent patterns. For English and Bai listeners, ratings for trials 1, 2, and 3 are relatively high, though Bai listeners rate trial 2 more highly than trials 1 and 3.

Among English listeners, ratings for trials 4, 5, and 6 are markedly lower than for trials 1, 2, and 3, with trial 5 receiving an especially low rating, at 3.10. For the English listeners, trial 6 stands in the middle of the range of mean ratings for this sub-category, at 3.90, while for the Bai listeners, this item is ranked the lowest among the constricted syllabic vocalizations, at 3.57. As discussed below, most of these differences appear to stem from degrees of laryngeal constriction in the utterances, combined with the extent to which a given utterance sustains the CV syllable patterns found in canonical and variegated babbling, versus the extent to which an utterance begins to drift into vocalic expression.

**Table 5-23. Ratings of constricted syllabic utterances, by trial and language.**

| Trial (Infant's Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| CS-1 (Bai) | 4.18 (1.299) | 4.18 (0.747) |
| CS-2 (Bai) | 4.53 (1.109) | 4.18 (0.984) |
| CS-3 (Bai) | 3.87 (1.436) | 4.15 (0.770) |
| CS-4 (English) | 4.00 (1.261) | 3.57 (1.059) |
| CS-5 (English) | 3.67 (1.366) | 3.10 (0.982) |
| CS-6 (English) | 3.57 (1.338) | 3.90 (1.057) |

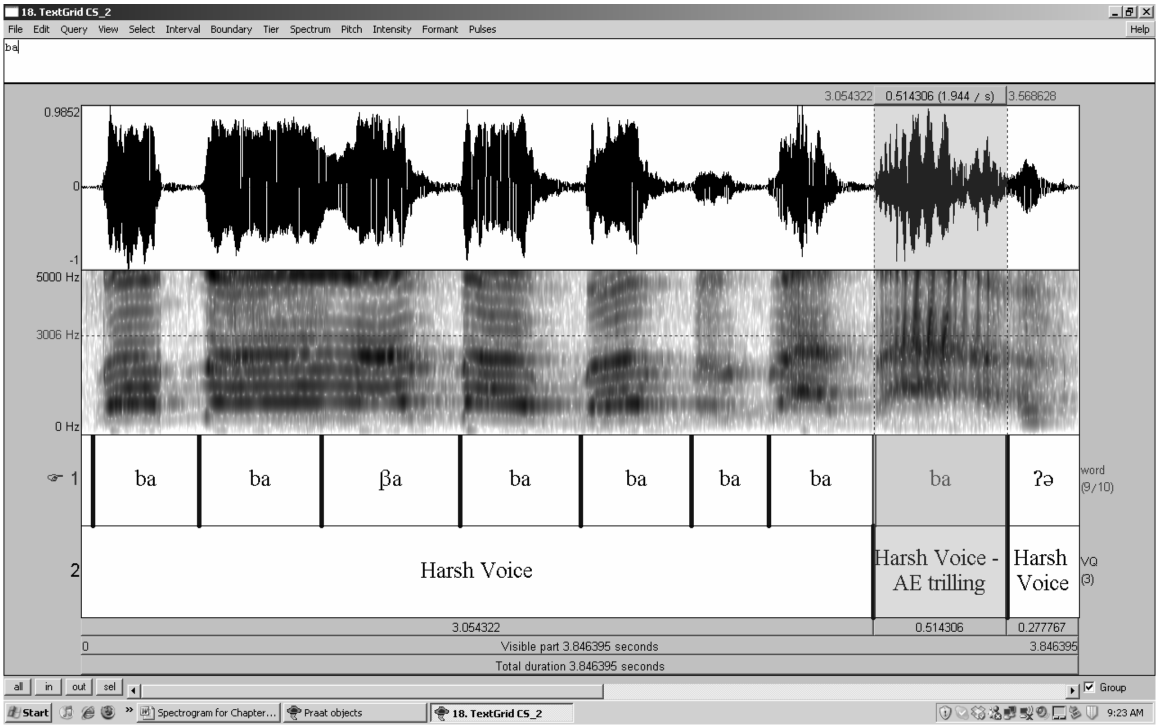
**Figure 5-28. English and Bai ratings of constricted syllabic utterances.**



Trial 1 is a voiced bilabial babbling sequence with central vowels, produced with mild laryngeal constriction throughout, and a falling contour. Like Trial 1, Trial 2 is a voiced bilabial babbling sequence with central vowels and a falling contour, but it is produced with very harsh voice in places, including aryepiglottic trilling towards the end of the utterance. Trial 3 is a voiced alveolar babbling sequence with central vowels, produced with moderately harsh voice and creaky voice at the end. Trial 4 is a variegated babbling sequence with voiced velar and alveolar nasal consonants and central vowels, produced with harsh voice throughout, but with very harsh voice towards the end of the utterance. Trial 5 is a voiced alveolar nasal babbling sequence, produced with harsh voice and, towards the end of the utterance, aryepiglottic trilling. During the period of aryepiglottic trilling, the infant has ceased producing CV syllables, and is producing harsh voice on a central voice with nasal resonance. Trial 6 is a variegated alveolar babbling sequence with very harsh voice, and a long central vowel in harsh voice at the end of the sequence; the syllables in the sequence vary in length and intensity, giving a greater impression of a word-like prosodic structure than seen in the other utterances.

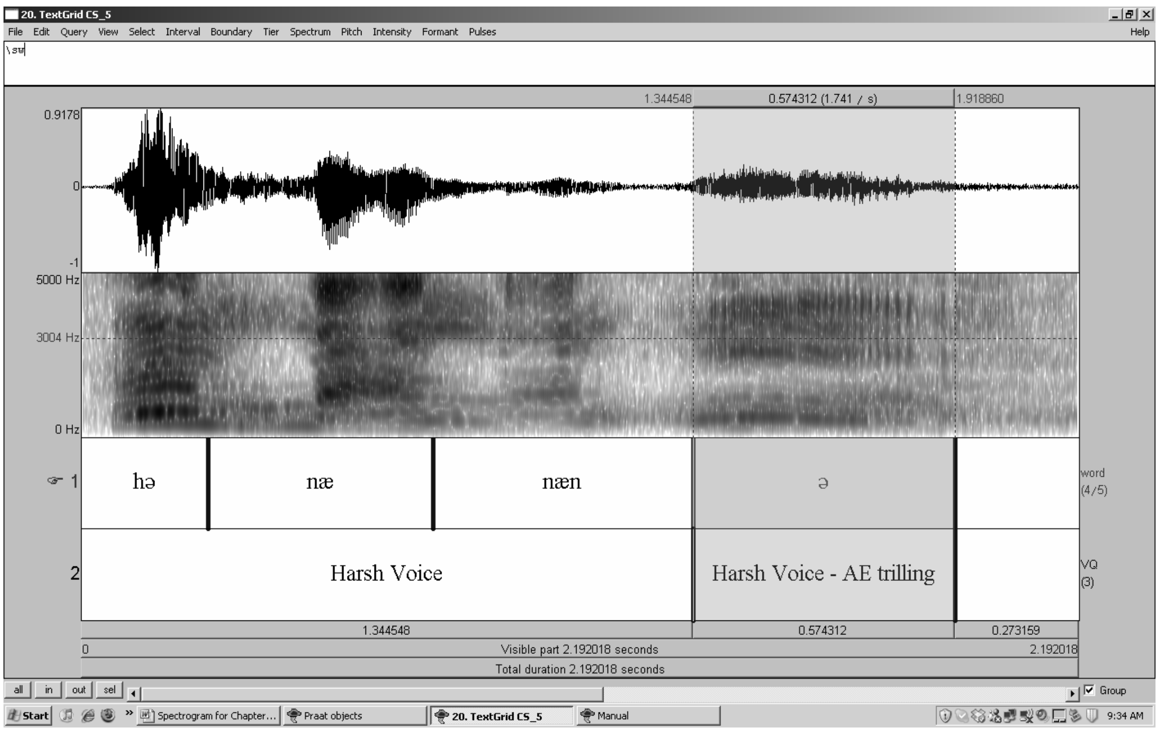
The constricted syllabic utterances that most closely correspond to canonical babbling receive high ratings from both language groups. However, Bai listeners give a particularly high rating to Trial 2, which mainly contrasts with Trials 1 and 3 in the stronger degree of harshness produced by the infant. Based on their nearly identical ratings of Trials 1, 2, and 3, it appears that English listeners are relatively insensitive to the degree of harshness in an utterance, provided a sequence of canonical syllables is maintained. By contrast, while Bai listeners prefer canonical babbling, the preference may be intensified by the presence of strong laryngeal constriction. Figure 5-29 below is a wide-band spectrogram of Trial 2. The aryepiglottic trilling is evident from the widely spaced striations in the upper part of the spectrum in the last syllable of the utterance.

**Figure 5-29. Spectrogram of constricted syllabic, trial 2.**




As noted above, ratings for trials 5, 6, and 7 decline among the English listeners, with the lowest ratings occurring for Trial 5. None of these utterances include as many syllables in a comparable time frame as Trials 1, 2, and 3. For example, Trial 2, shown above in Figure 5-29, is 2.64 sec in duration and contains seven syllables. Trial 5, shown below in Figure 5-30, is 2.24 sec in duration, and contains three syllables, followed by a vocalic sequence with aryepiglottic trilling with nasal resonance. The degree of harshness, and the presence of aryepiglottic trilling, is similar in trials 2 and 5. For the English listeners, the harshness can be disregarded in utterances with sustained syllabic production. However, once the utterance veers towards being non-syllabic, the degree of laryngeal constriction appears to weigh more heavily than it does in syllabic contexts in negatively influencing English listeners' ratings, while the same factor can influence Bai listeners in a positive direction. While Bai and English listeners both rate trial 5 lower than trial 2, the English rating is 3.10, the Bai rating is 3.67.

**Figure 5-30. Spectrogram of constricted syllabic, trial 5.**



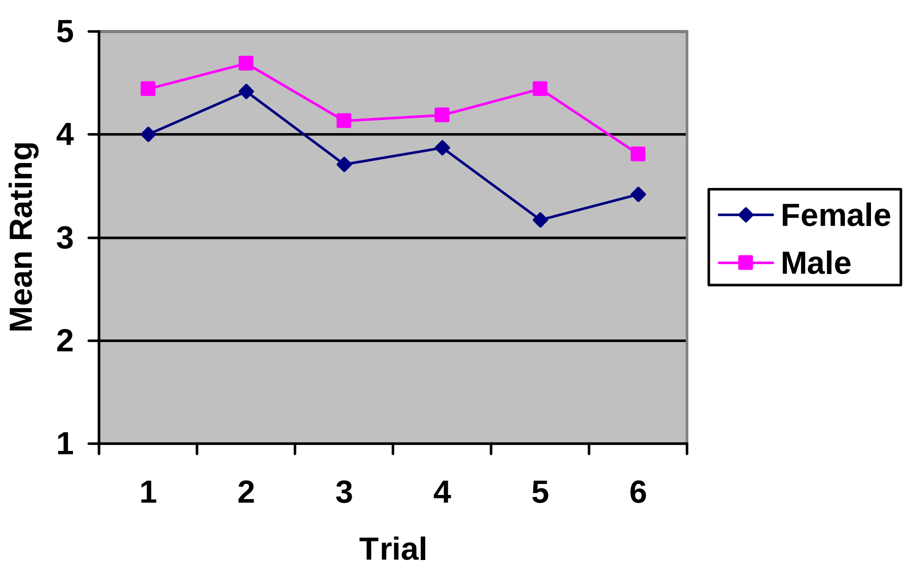

*5.4.1.2 Gender Differences*

As shown in Table 5-24 and Figure 5-31 below, Bai females rate all the constricted syllabic utterances lower than do the Bai males, a difference which is statistically significant ($F(1, 238) = 14.601$, $p > .005$). However, the relative ranking of individual items in the sample is the same across items for both genders, with the exception of trial 5, which is shown in Figure 5-30 above. It seems likely that female Bai listeners may be influenced by the same factor that was used to explain the low English ranking for this trial item: to the extent that an utterance contains non-syllabic portions, strong laryngeal constriction may negatively influence ratings. While the threshold for this factor may differ between the Bai female listeners and the English female and male listeners, with Bai females having a relatively higher threshold for laryngeal constriction, a similar effect may be in evidence. By contrast, Bai males may be biased in favour of laryngeal constriction, even when an utterance does not display a canonical syllabic pattern.

**Table 5-24. Bai ratings of constricted syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=24) | Male Rating (s.d.) (n=16) |
|---|---|---|
| CS-1 (Bai) | 4.00 (1.319) | 4.44 (1.263) |
| CS-2 (Bai) | 4.42 (1.283) | 4.69 (0.793) |
| CS-3 (Bai) | 3.71 (1.488) | 4.13 (1.360) |
| CS-4 (English) | 3.87 (1.329) | 4.19 (1.167) |
| CS-5 (English) | 3.17 (1.404) | 4.44 (0.892) |
| CS-6 (English) | 3.42 (1.283) | 3.81 (1.424) |

**Figure 5-31. Bai ratings of constricted syllabic utterances, by trial and gender.**

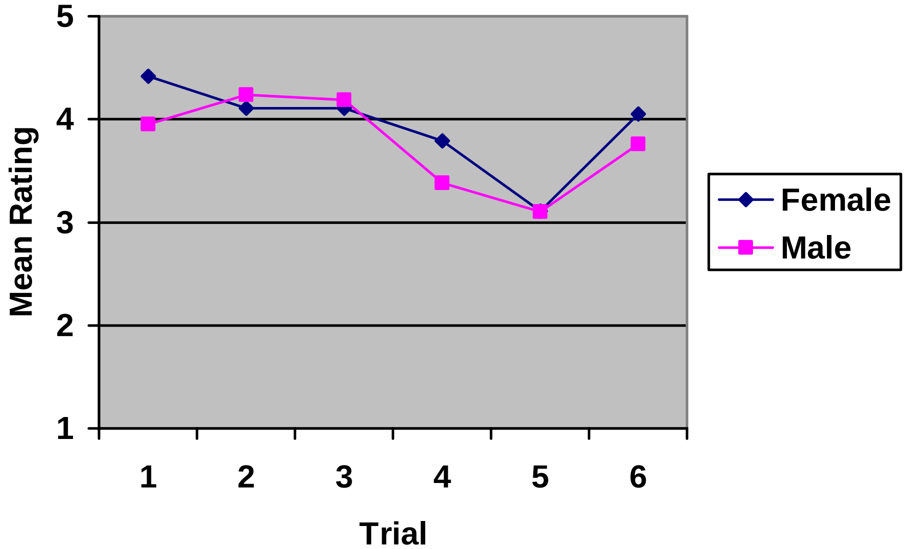


As detailed in Table 5-25, and as illustrated in Figure 5-32, the ratings of English females and English males do not differ significantly ($F(1, 238) = 1.064$, $p > .5$), beyond the fact that male ratings are noticeably lower than female ratings for trials 1, 4, and 6. There is no systematic reason for this difference, beyond the possibility that English males may rate laryngeal constriction somewhat lower than do women in syllabic contexts. However, if this were the case, one would expect male ratings for trial 2, the most harsh of the canonical babbling sequences, to be lower than male ratings for trials 1 and 3. In fact, however, English male listeners give this item the highest rating among constricted syllabic utterances. To date, I do not have a principled explanation for the slight differences in male and female ratings for these items.

**Table 5-25. English ratings of constricted syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| CS-1 (Bai) | 4.42 (0.607) | 3.95 (0.805) |
| CS-2 (Bai) | 4.11 (0.994) | 4.24 (0.995) |
| CS-3 (Bai) | 4.11 (0.875) | 4.19 (0.680) |
| CS-4 (English) | 3.79 (1.084) | 3.38 (1.024) |
| CS-5 (English) | 3.11 (1.100) | 3.10 (0.889) |
| CS-6 (English) | 4.05 (1.079) | 3.76 (1.044) |

**Figure 5-32. English ratings of constricted syllabic utterances, by trial and gender.**




### *5.4.2 Dynamic Syllabic*

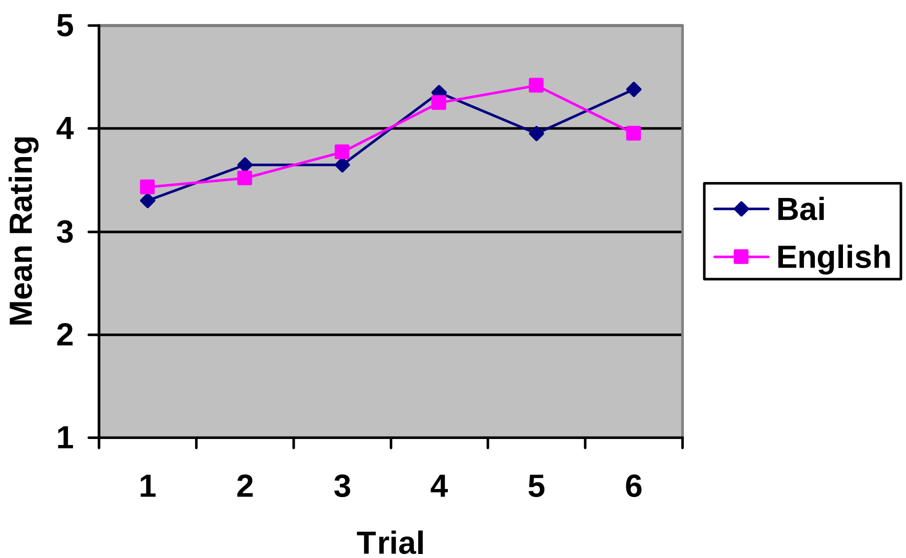
*5.4.2.1 Cross-Linguistic Comparison*

Bai and English mean ratings for dynamic syllabic utterances are similar in many respects, as reflected in the values reported in Table 5-26 and illustrated in Figure 5-33 below. Mean ratings for trials 1, 2, 3, and 4 are very similar. Also, ratings for trials 1, 2, and 3 are noticeably lower than trials 4, 5, and 6 for listeners from both language groups. Bai and English listeners differ in their relative ranking of trials 5 and 6. While English listeners assign a higher mean rating to trial 5 than to trial 6, Bai listeners show the reverse pattern. As outlined below, the shared patterns in the ratings of Bai and English listeners relate to the number of syllables included in the various utterances and to pitch range. The differences in the ratings between languages appear to stem from differences in the distribution of laryngeal voice quality variations over the utterance.

**Table 5-26. Ratings of dynamic syllabic utterances, by trial and language.**

| Trial (Infant's Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| DS-1 (Bai) | 3.30 (1.506) | 3.43 (1.130) |
| DS-2 (Bai) | 3.65 (1.578) | 3.52 (1.132) |
| DS-3 (Bai) | 3.65 (1.369) | 3.77 (0.947) |
| DS-4 (English) | 4.35 (1.210) | 4.25 (0.899) |
| DS-5 (English) | 3.95 (1.197) | 4.42 (0.675) |
| DS-6 (English) | 4.38 (1.030) | 3.95 (0.959) |

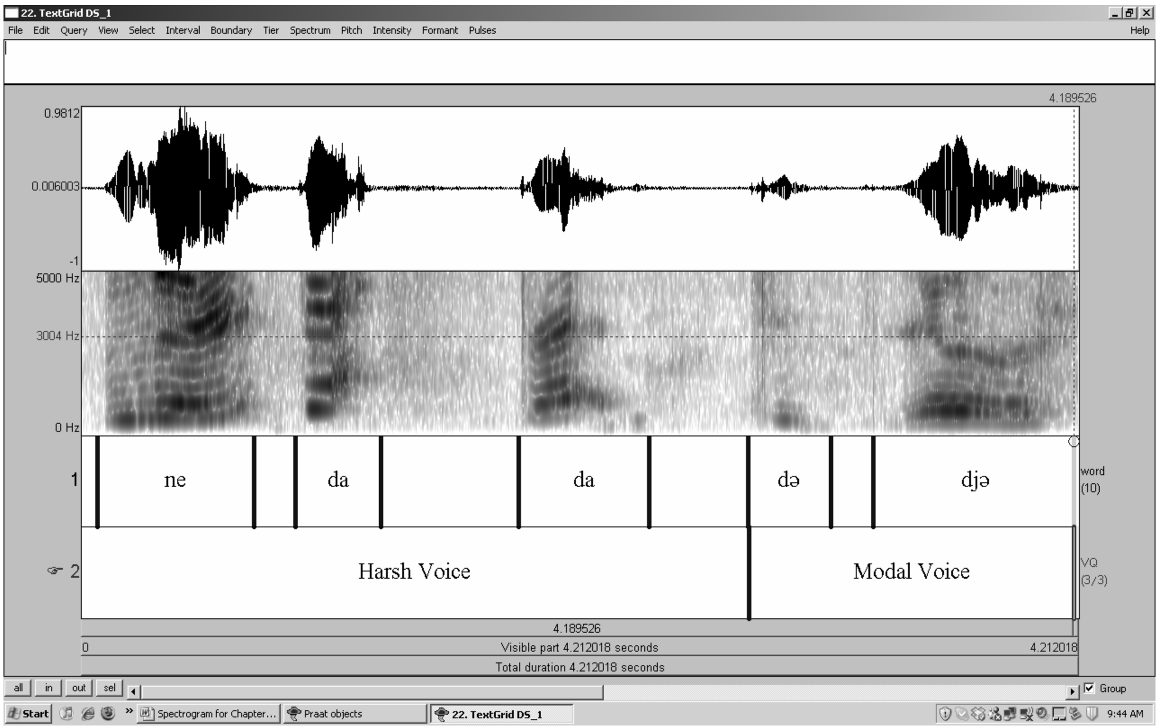
**Figure 5-33. Ratings of dynamic syllabic utterances, by trial and language.**




Trial 1 is a variegated alveolar babbling sequence with five slowly articulated syllables; the first three syllables are uttered in high-pitched harsh voice and the last two are produced with mid-pitched modal voice with a brief period of breathy voice at the end of the last syllable. Trial 2 is a variegated babbling sequence featuring palatal glides and voiced alveolar stops. The first five syllables in this sequence are produced with creaky voice and the last two syllables are uttered in modal voice. The auditory impression of trial 2 is one of monotony: the last five syllables in the sequence are produced with the same fundamental frequency. Trial 3 is a variegated babbling sequence including palatal glides and voiced alveolar stops. The nine-syllable sequence starts out in modal voice and then moves into a constricted setting with larynx raising, followed by whispery voice and finally trails off into low-amplitude, very low-pitched creaky voice. Trial 4 is a voiced nasal alveolar babbling sequence that alternates between modal voice and harsh voice over the nine syllables in the utterance. Trial 5 is a rapidly articulated series of five syllables. This voiced alveolar babbling sequence starts out in harsh voice and moves towards the middle of the utterance into modal voice, and then into falsetto at the end of the sequence. Trial 6 is a six-syllable babbling sequence that includes voiced bilabials and palatal glides. The sequence alternates between modal voice and harsh voice from syllable to syllable.

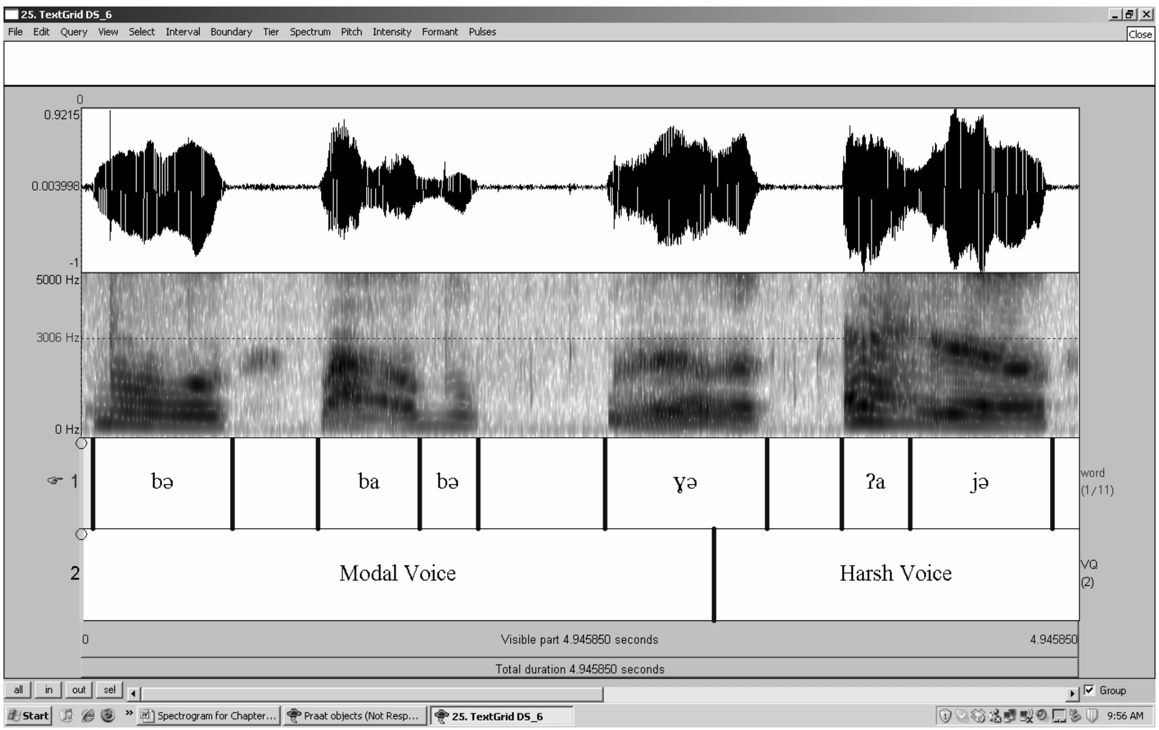
Across language groups, for trial 1, the relatively small number of slowly articulated syllables with comparatively long pauses in between seems to explain the low ratings assigned to this utterance, which is represented in Figure 5-34 below. However, the production of a long string of syllables is no guarantee of a high rating for either language group, as evidenced by the relatively low ratings for trials 2 and 3, compared to 4, 5, and 6. The narrow, monotonous pitch range of trial 2 and the slow trailing off of pitch and amplitude in the final syllables of trial 3, may be factors that influenced the ratings of these sounds for listeners in both language groups. Despite changes in voice quality in these utterances, both utterances convey an impression of “going nowhere,” prosodically speaking.

**Figure 5-34. Spectrogram of dynamic syllabic, trial 1.**




By contrast, those dynamic syllabic utterances that featured a sequence of clearly produced syllables within a moderate pitch range received higher ratings from both language groups, as reflected in the higher ratings of trials 4, 5, and 6 by Bai and English listeners alike. However, the language groups differed in their relative rankings of trials 5 and 6, and may have evaluated trial 4 using slightly different criteria, despite the highly similar mean rating on the latter trial (4.35 for the Bai listeners, and 4.25 for the English listeners). Based on the relative ranking of trials 4, 5, and 6, it appears that English listeners likely assigned trial 4 a relatively high ranking because it features nine rapidly articulated syllables. In turn, they may have assigned trial 6, depicted in Figure 5-35 below, a comparatively lower rating because it contains a smaller number of syllables that are produced quite slowly, though the pauses are not as long as seen in trial 1 (see Figure 5-34) above, and the changes in pitch and voice quality are not as marked as those seen in that utterance.

**Figure 5-35. Spectrogram of dynamic syllabic, trial 6.**




By contrast, Bai listeners may have assigned trial 4 a high rating not only because of the speed of articulation of the syllables, but because of the particular distribution of the changes in voice quality over the utterance. As noted above, trial 4 features a back-and-forth alternation between modal voice and harsh voice throughout the utterance. Bai listeners' preference for this type of alternation between constricted and unconstricted settings at the syllabic level is reflected in their relatively higher ranking of trial 6, shown in Figure 5-35 above, compared to English listeners. Like trial 4, trial 6 goes back and forth between harsh and modal voice over the course of the utterance. When English listeners are presented with dynamic syllabic utterances produced within a moderate pitch range by infants, they appear to give primary weight to factors such as the number of syllables and the speed of articulation—hence their preference for trial 6 over trial 4, despite the similarity in the distribution of the alternations in constriction in these two utterances. However, given a choice between two dynamic syllabic sequences that fit their criteria for syllables, they may prefer an utterance in which the alternation between different voice qualities is gradual, shifting from one quality into another, rather than going back and forth between settings from syllable to syllable—hence their preference for trial 5 over trial 4. Faced with these same choices, Bai listeners may give slightly greater weight to alternations in constriction from syllable to syllable—hence their preference for trials 4 and 6 over trial 5.

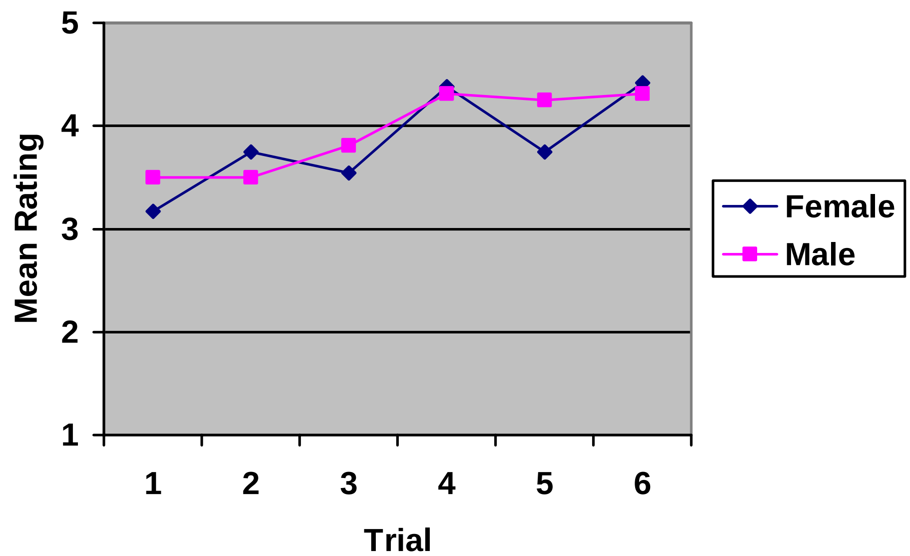
*5.4.2.2 Gender Differences*

The mean ratings of dynamic syllabic utterances by gender for Bai listeners are presented in Table 5-27 and Figure 5-36 below. There are no statistically significant gender differences for this sub-category ($F(1, 238) = .756$, $p > .5$). However, patterns in the relative rankings of individual utterances suggest that while Bai women and Bai men are influenced by many of the same factors in evaluating infant utterances, Bai women are more selective than Bai men in evaluating alternations in laryngeal constriction in syllabic utterances. The ratings of trials 4 and 6, which, as discussed above, feature a back-and-forth alternation of constricted and unconstricted voice qualities throughout utterances, receive the highest ratings among all the dynamic syllabic vocalizations included in this sample from Bai women and men alike. In addition, the mean ratings of Bai women and men nearly converge on trials 4 and 6, at 4.38 and 4.42 for Bai women, and 4.31 and 4.31 for Bai men.

**Table 5-27. Bai ratings of dynamic syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=24) | Male Rating (s.d.) (n=16) |
|---|---|---|
| DS-1 (Bai) | 3.17 (1.308) | 3.50 (1.789) |
| DS-2 (Bai) | 3.75 (1.539) | 3.50 (1.673) |
| DS-3 (Bai) | 3.54 (1.351) | 3.81 (1.424) |
| DS-4 (English) | 4.38 (1.209) | 4.31 (1.250) |
| DS-5 (English) | 3.75 (1.189) | 4.25 (1.183) |
| DS-6 (English) | 4.42 (0.974) | 4.31 (1.138) |

**Figure 5-36. Bai ratings of dynamic syllabic utterances, by trial and gender.**



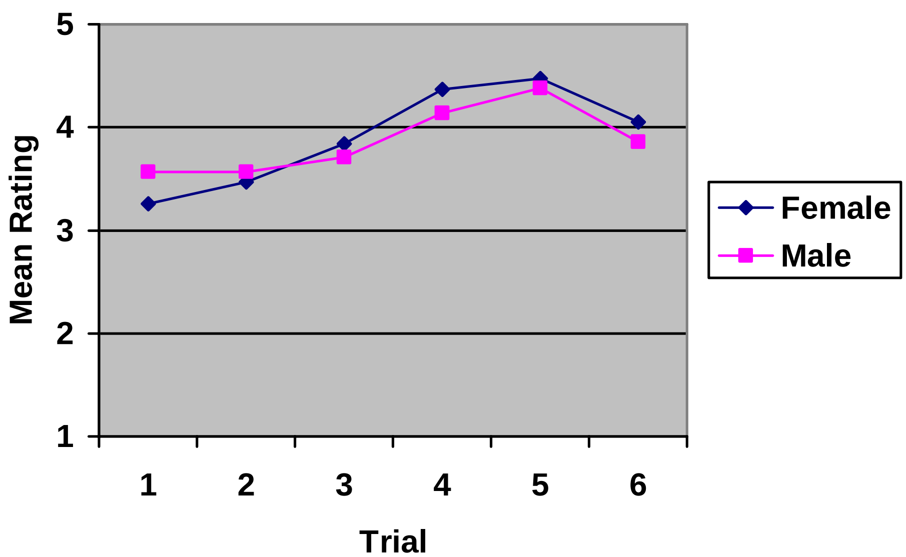

However, Bai women and men differ slightly in their ratings of trials 1, 2, 3, and 5. Bai men assign higher mean ratings to trials 1, 3, and 5 than do Bai women. In these instances, Bai men seem to be expressing an intrinsic preference for alternations in laryngeal quality, regardless of how they are distributed over the course of an utterance. By contrast, Bai women seem to be reserving their higher ratings for those dynamic syllabic utterances in which the distribution of alternations between constricted and unconstricted settings more closely corresponds to the distribution of such alternations within the Bai language, i.e., at the syllabic level. Bai males' general preference for variability in pitch and/or voice quality is also reflected in their assignment of a slightly lower rating to trial 2 than seen for the Bai women. As noted above, trial 2, while including variations in voice quality, stands out as being highly monotone in pitch among the dynamic syllabic vocalizations included in this perceptual task.

In keeping with the general pattern found in the English language group throughout this study, gender differences are not as noticeable as those sometimes found in the Bai language group, showing no statistically significant differences ($F(1, 238) = .008$, $p > 1$). As shown in Table 5-28 and Figure 5-37 below, the pattern of ratings for dynamic syllabic utterances is highly similar between English women and English men. In general, English women assign dynamic syllabic ratings slightly higher ratings than do English men, possibly reflecting a slightly greater preference for pitch variability within utterances relative to men, within a certain range, and provided the variability is smoothly distributed over the utterance. It is noteworthy that on trial 1, which features the most abrupt change in pitch and voice quality among the vocalizations in this sample, English women assign a lower rating than English men, at 3.26 versus 3.57, respectively. On trial 2, however, which is monotonous in pitch, English women award a very slightly lower rating than do English men, 3.47 versus 3.57.

**Table 5-28. English ratings of dynamic syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| DS-1 (Bai) | 3.26 (1.240) | 3.57 (1.028) |
| DS-2 (Bai) | 3.47 (1.124) | 3.57 (1.165) |
| DS-3 (Bai) | 3.84 (1.015) | 3.71 (0.902) |
| DS-4 (English) | 4.37 (0.761) | 4.14 (1.014) |
| DS-5 (English) | 4.47 (0.612) | 4.38 (0.740) |
| DS-6 (English) | 4.05 (1.129) | 3.86 (0.793) |

**Figure 5-37. English ratings of dynamic syllabic utterances, by trial and gender.**




### *5.4.3 Unconstricted Syllabic*

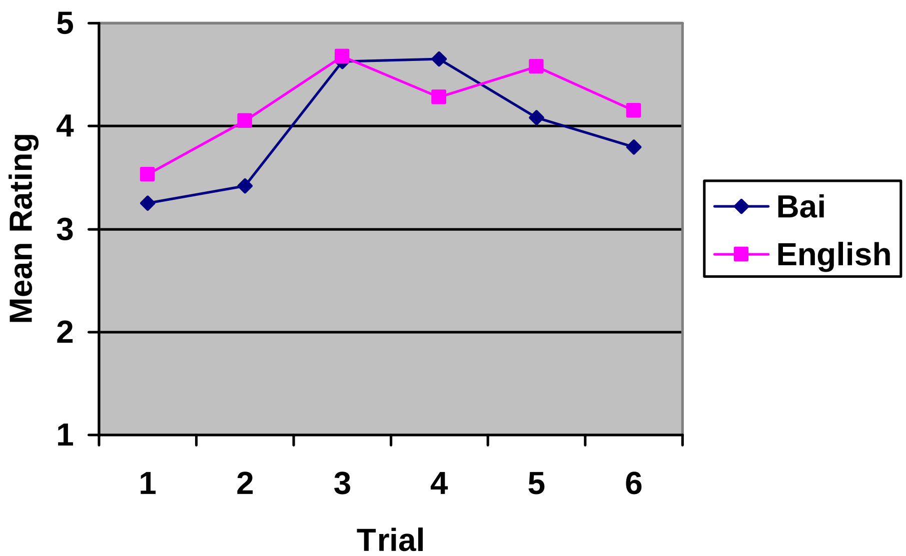
*5.4.3.1 Cross-Linguistic Comparison*

The mean ratings by trial and language are presented in Table 5-29. As depicted in Figure 5-38 below, most ratings for unconstricted syllabic utterances fall into a broadly similar pattern for participants in both language groups. Trials 1 and 2 receive lower mean ratings than trials 3, 4, 5, and 6 for Bai and English participants, and relative ratings for trials 1, 2, 3, 5, and 6 are the same for both groups. However, some differences are notable between the Bai and English participants. Mean English ratings for unconstricted syllabic utterances are higher for five of the six items within the sample, reflecting the tendency for English listeners to prefer unconstricted utterances. The only striking difference between the two language groups is that Bai listeners rate trial 4 higher than do English listeners. This difference may reflect an English preference for consistency in voice quality throughout an utterance, provided that an utterance satisfies other more highly ranked criteria that listeners apply to their evaluation of syllabic utterances produced by infants. These factors are discussed below.

**Table 5-29. Ratings of unconstricted syllabic utterances, by trial and language.**

| Trial (Infant's Language) | Bai Mean Rating (s.d.) (n=40) | English Mean Rating (s.d.) (n=40) |
|---|---|---|
| US-1 (Bai) | 3.25 (1.463) | 3.53 (0.987) |
| US-2 (Bai) | 3.42 (1.394) | 4.05 (0.904) |
| US-3 (Bai) | 4.63 (0.838) | 4.68 (0.694) |
| US-4 (English) | 4.65 (0.834) | 4.28 (0.751) |
| US-5 (English) | 4.08 (1.289) | 4.58 (0.594) |
| US-6 (English) | 3.80 (1.471) | 4.15 (0.893) |

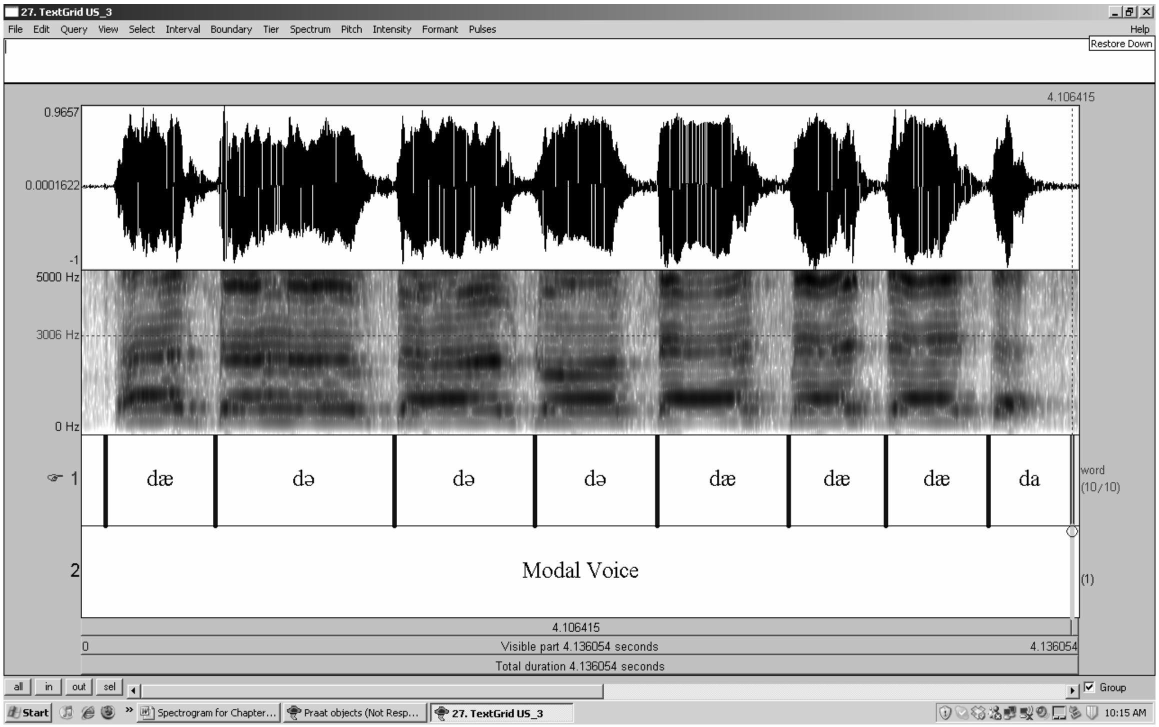
**Figure 5-38. Ratings of unconstricted syllabic utterances, by trial and language.**




Trial 1 is a variegated babbling sequence that includes voiced alveolar stops and nasals, and a palatal glide. The utterance begins with a vocalic portion in falsetto and then goes on in slightly lower-pitched, though still fairly high-pitched, modal voice for four syllables. Trial 2 is a babbling sequence of seven syllables, including voiced alveolar stops and a palatal glide, and modal voice throughout. The utterance is essentially a series of three shorter interconnected babbling sequences, the first comprising one syllable, and the second two consisting of two syllables each. Trial 3 is a variegated babbling sequence in modal voice with mostly voiced bilabial stops, but with two voiced alveolar stops in the two final syllables. This sequence is fairly level in voice quality and pitch, but features eight clearly and rapidly articulated syllables.  Trial 4 is a voiced bilabial babbling sequence which starts out with two syllables in breathy voice and then moves into a sequence of five rapidly articulated syllables in modal voice. Trial 5 is a variegated babbling sequence with six syllables produced with modal voice throughout. The consonants in the utterance are not fully formed stops, but rather, bilabial and velar approximants. Trial 6 is a variegated velar babbling sequence in which the first two syllables are produced in modal voice, and the last four in falsetto. As in trial 5, the velar consonants in this sequence are closer to being voiced approximants and fricatives than stops.

The preferences of the two language groups share many factors in common. Not surprisingly, the capacity to produce a series of rapidly articulated syllables in a unified sequence is an ability rated as highly important to language development among listeners from both language groups. The utterances in the sample that least reflect these characteristics, trials 1 and 2, receive the lowest ratings among unconstricted syllabic vocalizations from listeners of both language groups. Utterances that demonstrate this ability to a greater degree, i.e., trials 3, 4, 5, and 6, receive higher ratings from both language groups. The vocalization that most clearly demonstrates this capacity, trial 3, a spectrogram of which is shown in Figure 5-39 below, is among the two highest-rated utterances in the perceptual study among the English and Bai participants, at 4.68 and 4.63, respectively. It is also the utterance that shows the greatest convergence between the two language groups, not only in terms of mean rating, but also in terms of the standard deviation, which is 0.838 and 0.694 for the Bai and English listeners, respectively. In most other utterances in this study, the range of responses within the two languages tends to differ more than is the case with this utterance, with Bai listeners consistently showing a significantly larger standard deviation than English listeners.

**Figure 5-39. Spectrogram of unconstricted syllabic, trial 3.**

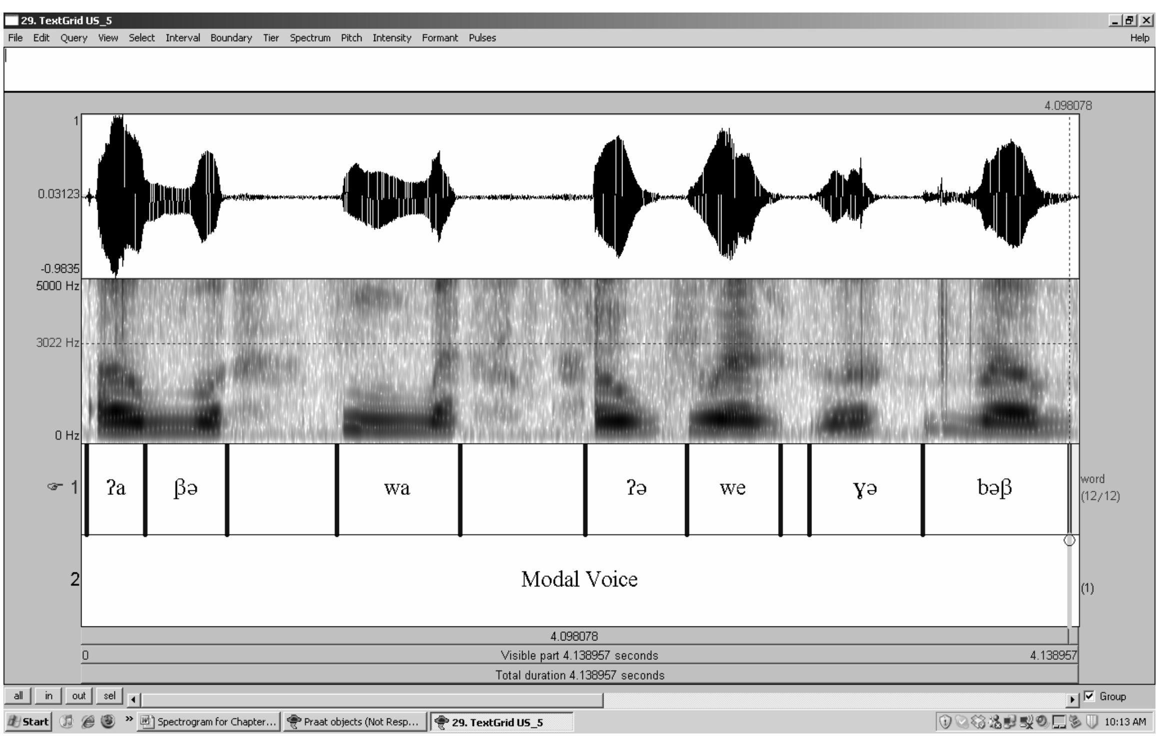



However, the ratings for unconstricted syllabic utterances also show interesting cross-linguistic differences. First, among the English listeners, trials 1 and 2, which are ranked lowest among unconstricted syllabic utterances by both language groups, are nonetheless rated more highly by English listeners than by Bai listeners, illustrating the English preference for unconstricted sounds. The English bias towards unconstricted utterances produced with stable voice quality, other things being equal, is also illustrated in the relative ranking of trials 4, 5, and 6 by the English and Bai participants. Trial 3, discussed above, is the highest-ranked utterance in the sample for the English participants, at 4.65, but trial 5 follows close behind, at 4.58, followed by trial 4, at 4.28. Among the Bai listeners, trial 3 and trial 4 receive virtually the same mean ranking, at

4.63 and 4.65, respectively, and trial 5 receives the lower mean rating of 4.08. Thus, English listeners slightly prefer trial 5 to trial 4, while Bai listeners prefer trial 4 to trial 5.

Among the utterances in this sub-category, trial 4, represented in Figure 5-40, is the most unconstricted of the vocalizations. Given infants' physiological predisposition to constriction, most sounds are produced with some degree of larynx raising, sometimes making the boundary between modal voice and mild laryngeal constriction hard to determine. Beyond a certain degree of larynx raising, the auditory and acoustic cues swing in favour of a perception of constriction; below that point, the sound is perceived as unconstricted. However, comparing the degree of energy in the upper portion of the spectrum in Figures 5-39 and 5-40, it seems clear that trial 4 is more unconstricted than trial 3. English listeners respond favourably to this characteristic, while Bai listeners are comparatively indifferent to it.

**Figure 5-40. Spectrogram of unconstricted syllabic, trial 4.**



Given a choice, English listeners also tend to prefer utterances produced with a single voice quality over utterances produced with an alternation in voice quality, even if that alternation takes place in an unconstricted laryngeal setting. By constrast, Bai listeners favour such alternations, other things being equal. This factor likely accounts for their preference for trial 4 over trial 5. Trial 5 includes an alternation between breathy voice and modal voice. Trial 6 also features an alternation in voice quality, from modal voice to falsetto. English listeners give lower ratings to trials 4 and 6 than to trial 5, which features modal voice throughout the utterance. On this basis, one might expect the Bai listeners' ratings to pattern with their mean rating for trial 4. However, trial 4 features more syllables, produced more quickly, than trial 6. Thus, in their relative evaluations of trials 4 and 6, as in their evaluation of other utterances in this and other sub-categories, Bai and English listeners may assign a different weight to a common set of criteria.

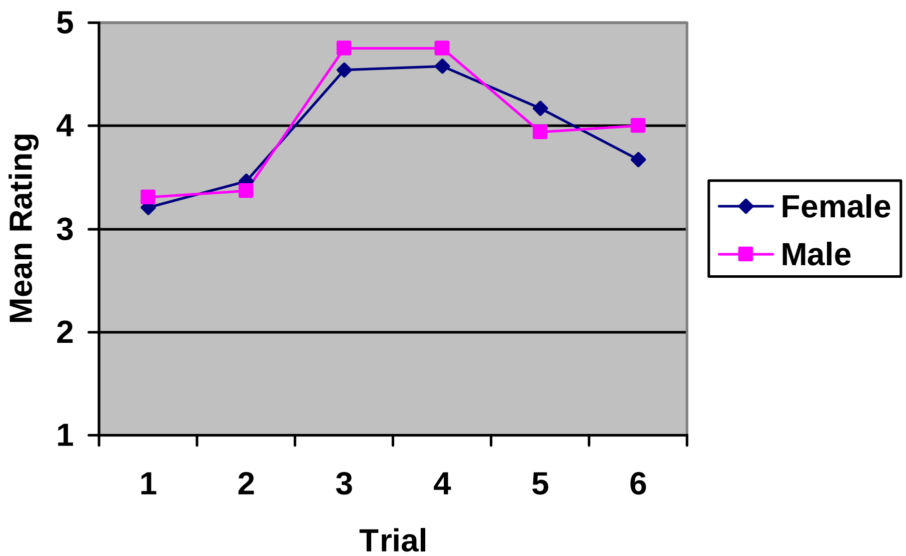
### *5.4.3.2 Gender Differences*

Table 5-30 and Figure 5-41 show the mean ratings of unconstricted syllabic sounds for Bai listeners, by gender. Bai men and women's relative ratings of individual utterances generally follow the same pattern, showing no statistically significant differences ($F(1, 238) = .400$, $p > 1$). Small differences in the ratings of individual items would appear to relate to a factor that has been identified in discussions of other sub-categories, i.e., Bai males' slightly greater preference for variability of pitch and/or voice quality in utterances of all types.

**Table 5-30. Bai ratings of unconstricted syllabic utterances, by trial and gender.**

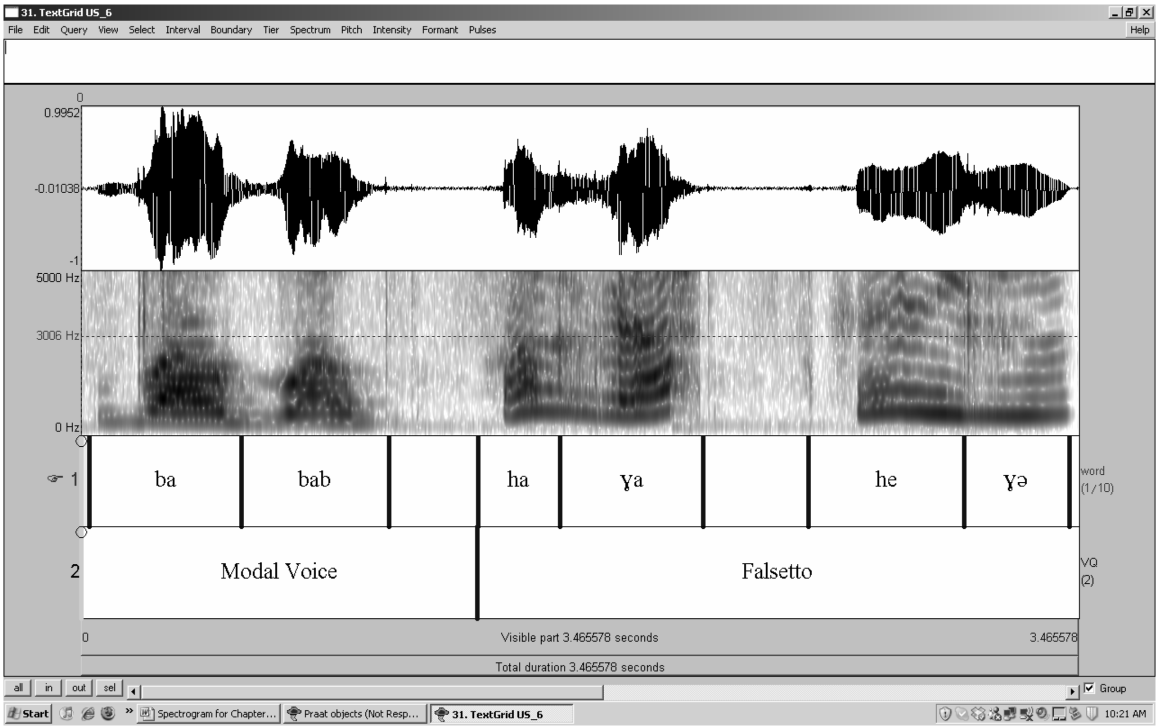
| Trial (Infant's Language) | Female Rating (s.d.) (n=24) | Male Rating (s.d.) (n=16) |
|---|---|---|
| US-1 (Bai) | 3.21 (1.351) | 3.31 (1.662) |
| US-2 (Bai) | 3.46 (1.285) | 3.37 (1.586) |
| US-3 (Bai) | 4.54 (0.977) | 4.75 (0.577) |
| US-4 (English) | 4.58 (0.881) | 4.75 (0.775) |
| US-5 (English) | 4.17 (1.090) | 3.94 (1.569) |
| US-6 (English) | 3.67 (1.551) | 4.00 (1.366) |

**Figure 5-41. Bai ratings of unconstricted syllabic utterances, by trial and gender.**




Trials 1, 4, and 6 include alternations in voice quality, though not of the type preferred by Bai women, who, as discussed in Section 5.3.2 above, show obvious preferences for voice quality alternations only when they occur from syllable to syllable. In the alternations seen in trials 1, 4, and 6, the infant produces a couple of syllables in one voice quality, and then a series of syllables in another quality. Figure 5-42, a spectrogram of trial 6, illustrates this pattern. As shown in the spectrogram, the first two syllables are produced in modal voice, while the last four are produced in falsetto. Bai men favour variability in general—hence their relatively higher mean ratings of trials 1, 4, and 6, compared to Bai women. In keeping with this pattern, Bai men assign a lower mean rating to trial 5 (see Figure 5-40 above), which is produced with modal voice throughout, than do Bai women.

**Figure 5-42. Spectrogram of unconstricted syllabic, trial 6.**

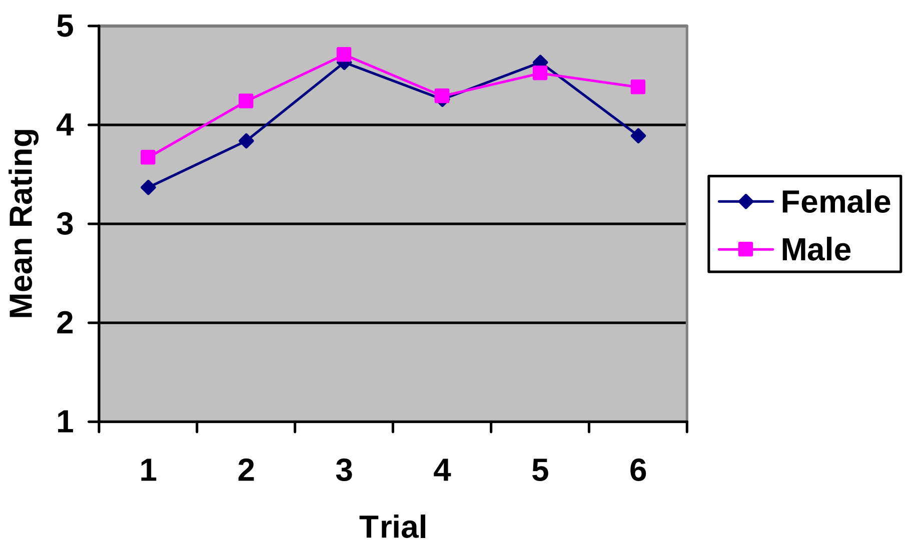



Mean English ratings for unconstricted syllabic sounds are shown in Table 5-31 and Figure 5-43 below. While the relative ranking of individual utterances within this sub-category is the same for both genders, compared to English female listeners, English male listeners tend to favour utterances on the basis of lack of constriction alone. This preference, which is statistically significant ($F(1, 238) = 4.639$, $p > .05$), is reflected in English males' higher mean ratings of trials 1, 2, and 6 relative to English females' mean ratings on these trials.

**Table 5-31. English ratings of unconstricted syllabic utterances, by trial and gender.**

| Trial (Infant's Language) | Female Rating (s.d.) (n=19) | Male Rating (s.d.) (n=21) |
|---|---|---|
| US-1 (Bai) | 3.37 (0.895) | 3.67 (1.065) |
| US-2 (Bai) | 3.84 (0.898) | 4.24 (0.889) |
| US-3 (Bai) | 4.63 (0.684) | 4.71 (0.717) |
| US-4 (English) | 4.26 (0.733) | 4.29 (0.784) |
| US-5 (English) | 4.63 (0.597) | 4.52 (0.602) |
| US-6 (English) | 3.89 (1.049) | 4.38 (0.669) |

**Figure 5-43. English ratings of unconstricted syllabic utterances, by trial and gender.**




### *5.4.4 Summary*

In this section, I discuss the findings for Bai and English listeners' ratings of syllabic utterances. These findings are summarized in Table 5-32 below.

**Table 5-32. Patterns in ratings of syllabic utterances, by language.**

| Sub-Category | Bai Patterns | English Patterns |
|---|---|---|
| Constricted Syllabic | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• Ratings for constricted syllabic are higher for Bai men than for Bai women.<br>• Ratings are positively affected by the presence of strong laryngeal constriction for Bai men, even when an utterance departs from a strictly canonical pattern. Bai women's ratings are negatively affected by the same factor. | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• Ratings are negatively affected by the presence of strong laryngeal constriction, especially when an utterance departs from a strictly canonical pattern. |
| Dynamic Syllabic | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• Ratings are negatively affected by monotonous pitch, particularly for Bai men.<br>• Ratings are positively affected by a larger number of alternations in voice quality at the syllabic level, particularly for Bai females. | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• English women rate dynamic syllabic utterances slightly higher than English men.<br>• Ratings are negatively affected by monotonous pitch, particularly for English women.<br>• Ratings are negatively affected by rapid changes in voice quality, particularly for English women.<br>• Ratings are positively affected if the utterance has a smaller number of |

| | | |
|---|---|---|
| | | alternations in voice quality. |
| Unconstricted Syllabic | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• Ratings are positively affected if the utterance alternates in voice quality, particularly for Bai men. | • Ratings are high for canonical babbling sequences featuring a series of rapidly articulated syllables.<br>• Ratings for unconstricted syllabic are higher than for other syllabic utterances, with the most unconstricted vocalizations receiving the highest ratings, particularly for English men.<br>• Ratings are positively affected when the utterance is produced with a single voice quality. |

Listeners from both language groups favour syllabic over non-syllabic utterances, reserving their highest ratings for syllabic vocalizations that feature a rapidly articulated series of syllables. The capacity to produce such sequences is important to speech and language development in all languages of the world, so this pattern is not surprising. There are, however, interesting differences in ratings between language groups, and between genders within language groups.

In general, Bai listeners do not distinguish among constricted, dynamic, and unconstricted syllabic utterances, though Bai females give slightly lower ratings to constricted syllabic utterances than to dynamic and unconstricted syllabic vocalizations. Across all sub-categories of syllabic vocalizations, Bai listeners show a preference for those utterances that feature alternations in pitch and/or laryngeal voice quality, compared to English listeners. Bai women are, however, more selective in their preferences in this respect. While Bai men express a general preference for variability in pitch and constriction, Bai women reserve their highest ratings for utterances in which pitch-dependent alternations in laryngeal voice quality are manifest at the syllabic level, as would be the case in the target language. This gender difference is consistent with one observed in the ratings of non-syllabic utterances. In making their ratings, Bai men seem to evaluate vocalizations on the basis of the presence or absence of a general feature, while Bai women are more sensitive to the manner and context in which that general feature is expressed. Thus, Bai women make a greater distinction between syllabic and non-syllabic utterances than do Bai men, and apply slightly different criteria in their evaluation of syllabic and non-syllabic utterances. For example, rapid alternations in pitch and laryngeal voice quality are characteristic of the Bai language, as is the presence of laryngeal constriction in general. Bai men respond favourably to these characteristics across all types of utterances presented in the listening task. While Bai women respond more favourably to these general features than do English listeners, they seem to make distinctions that Bai men do not. For instance, within non-syllabic contexts, Bai women are relatively biased against rapid changes in pitch and voice quality in favour of more gradual transitions. By contrast, in syllabic contexts, Bai women favour rapid transitions in laryngeal voice quality, while Bai men seem to respond favourably to the presence of transitions in voice quality regardless of how they are distributed over the utterance.

English listeners make a sharper distinction between syllabic and non-syllabic utterances than do Bai listeners. Within syllabic utterances, English listeners seem to respond primarily to the infant’s capacity to produce a series of rapidly articulated syllables. Once this criterion is met, their preferences for unconstricted laryngeal voice quality settings, and for utterances produced in a single voice quality, begin to influence their ratings. Compared to English men, English women favour utterances produced with variable pitch, provided such variations occur within a moderate range. While English women prefer pitch variations that occur in unconstricted utterances produced in a single voice quality, their relatively higher ratings on the dynamic syllabic utterances, compared to English men, likely reflect their general preference for pitch variation, even when this variation occurs in combination with changes in voice quality.

Overall, the ratings for syllabic and non-syllabic utterances show many common patterns, within and between languages, along with some general differences. However, ratings from both language groups vary widely, particularly for non-syllabic utterances. The wide variations in ratings are particularly striking among the Bai listeners. The results of the qualitative component of the study, discussed in Section 5.4 below, help to illuminate some of the factors that influenced listeners' ratings.

## 5.5 Qualitative Analysis

In Section 5.5, I discuss the findings of the qualitative component of the perceptual study, which is based on transcripts of brief interviews conducted with Bai and English participants following their completion of the listening task. As described in Chapter 3, in these interviews, participants were asked to reflect on the factors that influenced their ratings of the infant vocalizations. To review, participants were asked to rate the infant utterances on the basis of how important they perceived the presented sounds to be in learning to speak the target language (English or Bai). Listeners rated the sounds on a five-point scale, in which ‘1’ was intended to designate a sound that they considered not at all important in learning to speak the target language, and in which ‘5’ referred to a sound they considered to be highly important in learning to speak the target language.

Section 5.5.1 presents and discusses the results of the interviews with the Bai participants. In Section 5.5.2, I describe the comments made by the English participants.

### *5.5.1 Bai Interviews*

In this section, I outline the findings of the interviews with the Bai participants in the perceptual task. This discussion is organized into four broad sections. Section 5.4.1.1 describes participant comments that relate to laryngeal voice quality, and Section 5.4.1.2 outlines responses that relate to utterance type, though the two themes are not mutually exclusive. Where they overlap, the participants’ comments are included in the section considered most relevant to the discussion. Section 5.4.1.3 identifies other themes raised by the Bai participants in the interview. In Section 5.4.1.4, I suggest ways in which Bai participants’ comments in the interview relate to their performance on the listening task.

#### *5.5.1.1 Laryngeal Voice Quality*

The majority of the Bai participants explicitly demonstrated the role of laryngeal voice quality—and in particular, laryngeal constriction—in influencing their ratings of infant vocalizations in the perceptual study. While participants did not specifically refer to laryngeal constriction by name, 28 of the 40 Bai participants produced a sound that they said illustrated a property important in learning to speak Bai. Fifteen participants produced a non-syllabic constricted sound to illustrate the property they were referring to, usually a sequence of glottal stops followed by a vocalic portion that, according to this author's auditory judgement, featured varying degrees of laryngeal constriction. Degrees of harshness in these productions ranged from harsh voice with aryepiglottic trilling to a mildly constricted quality characteristic of larynx raising. A further 13 participants indirectly indicated the perceived importance of laryngeal constriction by producing a syllabic utterance (usually "bababa") that featured laryngeal constriction on at least one of the syllables.

Participants varied in their comments about the importance of laryngeal constriction in learning to speak Bai. Some participants, following their demonstration of laryngeal constriction, commented specifically on the importance of this type of utterance in learning Bai. In this vein, typical comments included: "most of Bai is like this"; "this is a very Bai-like sound"; "these are the most Bai-like sounds," and so on. One participant commented that such utterances "demonstrate proficient control of the Bai language." Some participants specifically referenced the way that adults' production might affect infants' production. One man commented that "we make this sound all the time, so the babies do, too." Another man noted the frequency of such utterances in the listening task, and commented that "it sounds like there are a lot of Bai babies in the experiment." One woman acknowledged that in making such sounds, babies may not be trying to speak Bai, but nonetheless sound Bai. In her words: "This is typical Bai-sounding speech. It might not have any particular meaning attached to it, but it sounds Bai." Finally, one man expressed ambiguity about whether the production of laryngeal constriction is specifically important in learning to speak Bai, or is of general importance in learning to speak any language. He said: “I’m not sure if this is important in learning to speak Bai, but it’s part of learning to speak, so I usually gave it a ‘3.’

*5.5.1.2 Utterance Type*

Many participants referenced the importance of syllabic utterances in learning to speak Bai. As with references to laryngeal constriction, these comments usually did not feature explicit mention of babbling or syllables, but rather, an imitation of a babbling sequence, most frequently, “bababa.” Twenty-five participants produced a babbling sequence as an example of the type of sound they heard in the listening task that they considered important in learning to speak Bai. Of these 25 people, 13 produced a babbling sequence that featured an alternation between a constricted and an unconstricted voice quality setting at the syllabic level. The remaining 12 people produced a babbling sequence in modal voice. One woman voiced the opinion that “when you hear ‘baba’ you know the sounds are developing into language.” On the other hand, some people commented that “it sounds like there are some babies who are learning Mandarin, and other who are learning Bai” and that “some of the babies don’t sound like they are learning Bai, but I can’t say why.” Possibly, these listeners were sensitive to different uses of pitch and constriction in the sounds, including the syllabic vocalizations.

Participants’ evaluations of non-syllabic vocalizations ranged considerably. It appears that some participants considered constricted non-syllabic vocalizations to be important in learning Bai, as discussed above in Section 5.4.1.1. Others, most likely referring to constricted and/or dynamic non-syllabic vocalizations, questioned the relevance of these sounds in learning to speak Bai. For example, one man said that “some of the sounds are just screaming and shouting, and are not speech-like at all.” In a similar vein, another man reported that "some of the sounds are like speech, but others are just animal imitations." One man, following his imitation of a constricted non-syllabic vocalization, said that "the baby is just crying out there, not learning to talk." However, this same man, moments before, stated that such sounds were "prototypical Bai sounds." One woman expressed uncertainty about how to evaluate some of the non-syllabic sounds. As she said, "sometimes the baby is just screaming, but this can still be an important part of learning to talk."

Extra-linguistic factors pertaining to communication and the expression of emotion also figured in people's evaluations of non-syllabic sounds. One woman said that "it sounds like some of the babies are sometimes uncomfortable or tired." Two people mentioned that, in the words of one woman, "it sounds like the baby is asking for something."

#### *5.5.1.3 Other Factors*

Listeners reported that their responses were affected by other extra-linguistic factors, including their own experiences of children, and their knowledge of developmental processes. One woman said: "It's hard to describe … I just kept comparing the sounds to what I remember my own child doing when he was a baby." An older man said that "a lot of the sounds reminded me of my grandson, who is learning to talk right now." Sounds that reminded the listeners of a child they know and care for may sometimes, thus, have generated higher ratings, independent of the listeners' judgment of the importance of the sound in learning to speak Bai. On the other hand, knowledge and experience of stages that babies go through may have prompted some listeners to give lower ratings to non-syllabic sounds in general. One woman said that "it sounds like there are some younger babies in there," echoing the comments of another woman, who shared her impression that "some of the sounds are made by younger babies, and some of them are made by more advanced babies." Such impressions may have caused some people to give lower ratings to sounds that they know appear earlier in language development, even if they were aware of the importance of those sounds in Bai.

*5.5.1.4 Discussion*

Based on the comments shared by participants following the listening task, it appears that most listeners consciously considered laryngeal constriction and utterance type in evaluating the infant vocalizations, even if they did not use phonetic terms to describe these features. Perhaps the most noteworthy contribution of the interviews, in terms of interpreting the results of the perceptual task, lies in highlighting the conflict that many listeners felt in evaluating constricted and/or dynamic infant vocalizations, particularly non-syllabic constricted and dynamic utterances. In the interviews, some listeners clearly stated the importance of such sounds in learning to speak Bai, while others (12 of the 40) did not mention this factor and expressed only a preference for syllabic sounds. Another group expressed uncertainty in evaluating the significance of constricted and/or dynamic vocalizations: while they were aware of the importance of the laryngeal voice quality features contained in these sounds in Bai, they were also aware that infants likely were not explicitly trying to speak Bai in producing these sounds; that infants know how to produce such sounds early in life, before language development is in full swing; and that infants often use such sounds for other purposes, such as expressing emotion and making requests. Thus, listeners may have felt pulled in several directions in rating these sounds, possibly accounting for the large standard deviations among Bai listeners, particularly for non-syllabic utterances.

### *5.5.2 English Interviews*

Generally, English participants were more voluble in expressing their opinions about the infant sounds they heard in the study. While nearly all of the Bai participants' comments are reported in Section 5.5.1 above, only a small proportion of the comments offered by English participants will be included here, with the intention of highlighting general trends.

The discussion of the English interviews is organized in the same manner as that employed for the discussion of the Bai interviews. Section 5.5.2.1 discusses English participants' comments about laryngeal voice quality, while Section 5.5.2.2 highlights their perspectives on utterance type. Once again, the discussion of laryngeal quality is closely connected with the discussion of utterance type, so some comments about laryngeal quality will be included in the discussion of utterance type, when a participant's comment focuses specifically on the relationship of voice quality and utterance type. Section 5.5.2.3 focuses on other factors mentioned by English participants as affecting their ratings. Finally, Section 5.5.2.4 summarizes the findings of the interviews and discusses how the main trends in the interviews may relate to the English participants' performance on the perceptual task in this study.

*5.5.2.1 Laryngeal Voice Quality*

The majority of English participants—28 of 40—referred to the strong influence of laryngeal voice quality features on their ratings of the infant vocalizations, though most people did not have the knowledge or vocabulary to distinguish voice quality features from inter-related prosodic features, such as pitch and loudness. In pointing to the influence of laryngeal voice quality features, English participants occasionally referred to "guttural, throaty sounds, " but generally used a range of impressionistic terms to try to convey their preferences and aversions, including soft versus hard, female versus male, smooth versus rough, controlled versus uncontrolled, moderate versus extreme, rational versus emotional, happy versus upset, and so on. Generally speaking, English listeners expressed a preference for the first term in each of the latter contrasting sets, which often seemed to refer to unconstricted versus constricted sounds, respectively, though some variations in preferences were expressed.

For the most part, English participants' comments reflected an aversion to laryngeal constriction. One woman's comment reflects the common theme expressed by most participants: "The loud, harsh sounds were less speech-like, and the softer ones were more speech-like, rather than just an accidental gush of energy. The soft ones seemed like the child was able to form something." One man reflected that he "preferred the ones without 'sound effects' like growling, shrieking, and shrilling." Similarly, one woman reported that she "didn't like the screaming, high-pitched, aggressive sounds" and that she "preferred the lower-pitched, smoother sounds," which seemed to her "more controlled." This perspective was echoed in many participants' comments, only a few of which are reported here.

While, as stated above, most participants expressed an aversion to laryngeal constriction, some people noted its potential communicative efficacy. One English woman, who reported giving high ratings to constricted sounds, said: “When I heard the harsh and loud sounds, I knew the baby had some intention to communicate, which is different from what they do when they’re just trying to entertain themselves. I was biased towards these sounds, because this is what learning to talk is all about: pairing sound with meaning.” By contrast, some participants, while recognizing infants’ use of constricted and/or dynamic sounds in gaining adults’ attention, clearly wanted to encourage an entirely different pairing of sound with meaning. As one man put it, “with our son, there were certain sounds that we tended to reinforce or discourage. We tried to ignore loud requests and respond to soft ones.” A similar, perhaps unintentionally comic, perspective on encouraging culturally appropriate use of the voice in communication is reflected in this woman’s comment that “in English, we don’t screech at each other—there’s an emotion attached to that.”

Indeed, the question of what infants are doing when they produce constricted and/or dynamic sounds, particularly non-syllabic utterances, is one that many English listeners struggled with in evaluating the sounds in the perceptual test. Listeners expressed various opinions on whether constricted and/or dynamic utterances were communicative resources, expressions of emotion, and/or examples of infants playing with their voices. These reflections seemed connected to a general difficulty in distinguishing speech, language, and communication, and their relationship to laryngeal voice quality and other prosodic features, compounded by cultural understandings of the appropriateness of various forms of expression. Some evocative comments in this vein include one woman's comment that she "gave higher ratings to the ones where communication started to enter into it—language is not just noise" and one man's reflection that "I realized that I was beginning to judge by emotion—if it sounded really harsh, then it seemed like an expression of emotion, rather than an ability to communicate. The quieter sounds seemed more deliberate, like there was more of an attempt." This woman's statement is a further example of the issues that English listeners grappled with in making judgments on the sounds:

> I had a hard time distinguishing emotional expression from 'speech necessities,' but when the emotion seemed less extreme, I thought it was more like speech, and when it sounded like the baby was 'playing with sounds' rather than expressing pure emotion, I gave it a higher rating. But expressiveness can't be separated from speech and language, because speech is designed to carry emotion.

Some listeners reported making a deliberate effort to ignore constriction in evaluating the sounds, seeming to reflect an intuitive awareness that laryngeal voice quality differences are not contrastive in English, and therefore, should not be allowed to influence their judgments of whether the sounds were important for infants in learning to speak the language. As one woman put it, "the sounds that were clearer with less expression and less emotion were probably more speech-like, and so at first I gave them higher ratings, but then I began to correct for this, because it was biasing me." Echoing this same perspective, another English woman said: "I tried to compare the sounds in their own right and ignore the emotional difference, even though some sounded more 'clear' to me than others—I tried to think of them as the same." Finally, in commenting on her criteria in rating the sounds, one woman stated:

> I tried not to pay attention to the loudness and the pitch. Sometimes I gave the back, guttural sounds lower ratings than the softer ones at the front of the mouth, but after a while, I tried to ignore this and think of the sounds as the same, even though I could clearly hear the differences.

*5.5.2.2 Utterance Type*

English participants' ratings in the perceptual task reflected a clear preference for syllabic utterances. In the interviews, 29 people cited their preference for babbling sequences, sometimes, like the Bai participants, by producing imitations of babbled utterances, and at other times, by mentioning their preference for consonants in infants' vocalizations. As one man stated, "the ones with lots of syllables are more focused toward language—in those sounds, the tongue is in control." One man said that "the more consonants I heard, the higher the rating I gave the sound." This perspective was echoed in the comments of many other participants, who made statements to the effect that the production of consonants involved more vocal motor control, more breath control, and greater ability to "shape the sound" as opposed to "just opening their mouth and letting it all hang out."

Some listeners commented on their attitudes towards laryngeal voice quality variations in babbling, with the greater number expressing a preference for unconstricted babbling. One woman reported that she "went for the more conversational sounds. Some of the sounds were like speech, but were so high-pitched—when they were lower-pitched, they sounded more like a rational speech rhythm, but of course, in real life, if the babies were agitated and needed something, they'd make these higher sounds." One man admitted that he "was a sucker for the soft 'bababababa' sounds. I didn't like the guttural sounds, even though they're language, too." The occasional listener reported trying to be accepting of the constricted and/or dynamic syllabic sounds. One woman noted: “I was looking for more mature-sounding speech, but I also tried to embrace the extremes, because all these variations are present in various contexts in English, even though you don’t hear them on the 6:00 news.”

#### *5.5.2.3 Other Factors*

Like some of the Bai participants, many English participants referred to their own experiences with children and/or their knowledge of stages in speech and language development as factors that influenced their ratings. One woman said that she “was trying to picture the age of the child—they didn’t seem to all be from the same age group, and this affected my judgment.” Another man commented that “some of the sounds seem to be made by a ‘wee babe’—a much younger baby than the other sounds.” While some participants noted a tendency to assign lower ratings to sounds they thought of as “earlier,” other participants cited developmental knowledge and/or parenting experience as positively influencing their ratings of “less advanced” sounds. As one woman noted, “even though the early sounds were kind of shrill, I remember those sounds and I think they’re still like the baby trying to learn to form words and talk.” In a similar vein, one older woman commented that “I’ve been around a lot of babies and I can hear ‘talking’ in almost anything. Babies try so hard—they’re always trying.”

#### *5.5.2.4 Discussion*

The comments offered by the English participants are consistent with many of the results reported for the perceptual experiment. While some English adults voiced an appreciation for the potentially important role of dynamic and/or constricted sounds in communication, emotional expression, and speech development, and while a small number of others reported trying to ignore laryngeal voice quality variations, most adults stated a clear preference for unconstricted sounds, whatever the terms they used to describe such sounds. This preference is reflected in the higher mean ratings for unconstricted utterances across syllabic and non-syllabic vocalizations.

However, English adults' conflicting views on the relevance of laryngeal voice quality variations are also clear in the results of the perceptual test, particularly in the ratings for non-syllabic vocalizations, most of which have larger standard deviations than the ratings for syllabic utterances. English listeners express a strong preference for syllabic over non-syllabic utterances. While English attitudes towards laryngeal voice quality variations are also reflected in their ratings for these items, the variation takes place within the context of much higher mean ratings, and hence, spans a narrower range, as reflected in the smaller standard deviations in rankings for most syllabic utterances.

## 5.6 Summary

In Section 5.5, I summarized the comments made by Bai and English participants following their completion of the perceptual task in attempting to account for factors that influenced their ratings of the infant utterances, as reported in Sections 5.1, 5.2, and 5.3, and 5.4. Both language groups express a preference for syllabic utterances, which is consistent with the higher mean ratings assigned to such utterances by Bai and English listeners in the listening task. On the question of laryngeal voice quality, both groups are pulled in more than one direction, though the conflict is greater for the Bai participants than for the English participants. English participants face the challenge of evaluating the purpose, relevance, and/or appropriateness of laryngeal voice quality variations in infant vocalizations, within the context of their knowledge of developmental processes and a cultural preference for unconstricted voice quality settings. Like the English participants, Bai adults also consider developmental and extra-linguistic factors in evaluating laryngeal voice quality variations, but these judgments are complicated by the fact that their language makes extensive use of such variations. Thus, while listeners from both groups are pulled in sometimes conflicting directions in evaluating laryngeal voice quality variations in infant utterances, the Bai listeners are pulled in more directions, and to a greater extent, than are the English listeners. The directions in which adults listeners are pulled may be systematically related to the patterns in infants' production in the two language groups. This relationship will be explored more fully in Chapter 6.

# Chapter 6

# DISCUSSION

In this chapter, I discuss the results of the findings of the production and perception components of the present study in relation to the research questions outlined in Chapter 1, and in relation to the literature on production and perception reported in Chapter 2. Section 6.1 discusses the results of the production study, and Section 6.2 discusses the findings of the perceptual task, along with the relationship between the patterns observed in infants' production and adults' perception.

## 6.1 Production Results

The discussion of the production study, while not explicitly organized around the research questions outlined in Chapter 1, is intended to address the main themes mentioned in those questions. These questions are repeated below from Chapter 1, for convenience:

- **Question 1:** How does the use of laryngeal voice quality features, particularly the use of laryngeal constriction, change over the first year of life?
- **Question 2:** Does the incidence of laryngeal constriction differ according to the type of utterance that infants produce? In particular, does the use of laryngeal

constriction differ in syllabic and non-syllabic utterances in general and over time?

- **Question 3:** What is the relationship between the development of laryngeal control and the emergence of control within the oral vocal tract?
- **Question 4:** Are there universal patterns in the development and use of laryngeal voice quality features in the first year of life?
- **Question 5:** Does infants' use of laryngeal voice quality features differ according to the use of these features in the infants' ambient language? If so, do these differences affect other features of infant speech development, such as the development of babbling and the development of control over the oral vocal tract?

This discussion highlights those aspects of the present study that are most relevant to the research questions above, and that address gaps in the existing literature on infant speech production. Sections 6.1.1 and 6.1.2 discuss universal and language-specific patterns in the production of laryngeal voice quality and utterance type, respectively. Section 6.1.3 focuses on the distribution of laryngeal voice quality in mixed and syllabic utterances, a discussion that highlights language-specific patterns in the development and integration of control over the laryngeal and oral vocal tracts. Throughout these sections, several figures that were introduced in Chapter 4 are reproduced for the reader's convenience.

### *6.1.1 Laryngeal Voice Quality*

The present study, while exploratory in nature, provides the most extensive phonetically grounded analysis of laryngeal constriction in the first year of life currently available in the literature. In previous studies of infant speech production, researchers have noted the presence of laryngeal constriction in the early utterances of infants from a range of Indo-European language backgrounds (Koopmans van Beinum & van der Stelt, 1986; Nathani et al., 2006; Oller, 1980, 2000; Roug et al., 1989; Stark et al., 1975). However, these researchers have described these sounds impressionistically, limiting opportunities to systematically document patterns in the development of laryngeal constriction in the first year of life, and obscuring the role of laryngeal constriction in phonetic development as a unitary factor. In one of the first phonetically grounded studies of laryngeal voice quality in early infancy, Bettany (2004) noted the prevalence of such sounds in the first six months of life in her observations of one English infant. Using a similar phonetic framework to that employed by Bettany, related findings were reported by members of the InSpA team at the University of Victoria, in their preliminary analysis of English, Bai, and Moroccan Arabic infants in the first year of life (Benner et al., 2007; Esling et al., 2004a, 2004b, 2004c, 2006). The results of this study confirm and extend Bettany's (2004) findings to a larger group of infants, from two language backgrounds, over the whole of the first year, and provide a systematic consolidation of the preliminary results reported by the InSpA team for English and Bai infants.

As noted in Chapter 2, infants begin their lives with a physiological predisposition to produce laryngeally constricted sounds (Crelin, 1973; Eckel et al., 1999; Fried et al., 1982; Kent & Vorperian, 1995; Sasaki et al., 1977; Vorperian et al., 1999). Consistent with this physiological setting, the vast majority of sounds produced by Bai and English infants in months 1-3 are marked by laryngeal constriction, as illustrated in Figures 6-1 and 6-2, below. However, as noted in the medical literature cited above, starting in the third month of life, the infant vocal tract undergoes significant growth and restructuring. The larynx, epiglottis, and hyoid bone begin to descend, lengthening the vocal tract and changing the angle of the oral cavity relative to the pharyngeal cavity, increasing the range of pitch, intensity, and constriction that infants may produce. In keeping with this pattern of physiological development, the infants in this study begin to produce alternations between constricted and unconstricted voice quality settings in increasing numbers in months 4-6, as reflected in the significant increase in dynamic utterances during this period for infants from both language groups. Months 4-6 have been noted by other researchers as a period in which most infants begin to engage in “vocal play” or “vocal entertainment” (Oller, 2000). The findings of this study clarify that for infants from both language groups, vocal play involves, among other things, a systematic alternation between laryngeal voice quality settings, and that systematic exploration of these settings continues throughout the year, as reflected in the continued strong presence of dynamic utterances during the first year for Bai and English infants. Infants from both language groups also produce an increasing proportion of unconstricted utterances throughout the year.

**Figure 6-1. Laryngeal quality in Bai utterances, by age and percentage.**

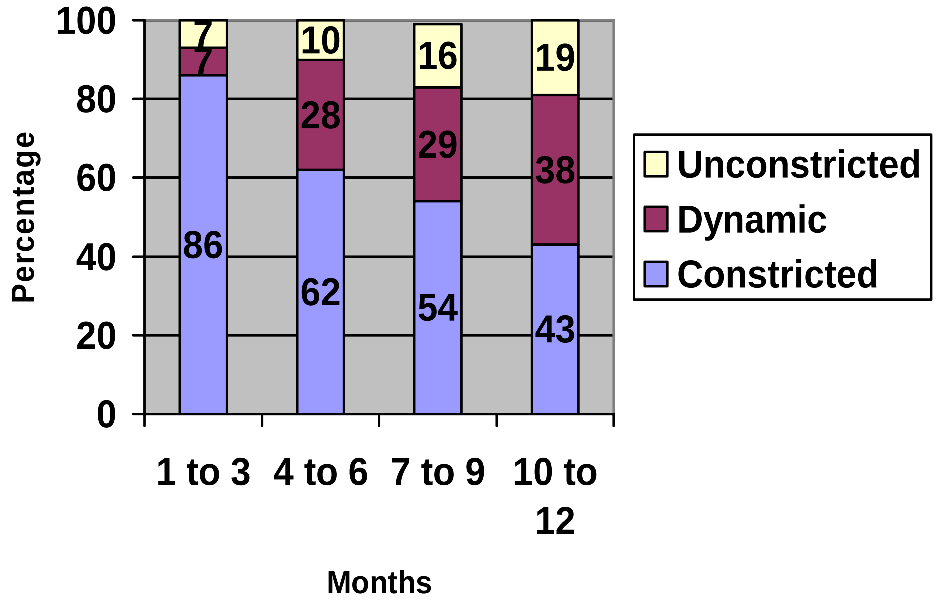
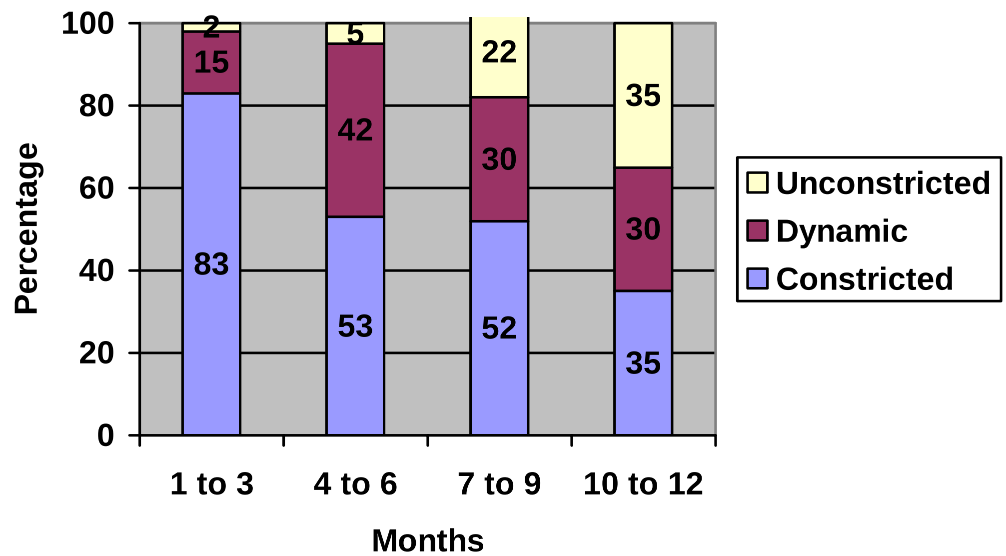



**Figure 6-2. Laryngeal quality in English utterances, by age and percentage.**




The overall patterns in the development of laryngeal voice quality likely reflect universal trends, but there is no basis upon which to definitively make such a statement, because no other systematic studies of the development of laryngeal voice quality currently exist in the infant speech literature for infants of any language background. The model employed in this study would need to be extended to studies of infants from other language groups to confirm whether the patterns documented herein are in fact universal.

Differences in the development of laryngeal constriction between the two language groups likely reflect the influence of the ambient language. Constricted utterances decline more slowly, and to a lesser extent, among Bai infants than among English infants. Dynamic utterances increase in frequency throughout the year among Bai infants, but decline somewhat for English infants. Conversely, unconstricted utterances increase more dramatically in the second half of the year among English infants compared to Bai infants. These overall patterns suggest that towards the end of the first year, the production of laryngeal voice quality is beginning to reflect the use of such features in the target language. However, these findings need to be confirmed in studies of more Bai and English infants, and in studies of infants from other language backgrounds that employ laryngeal voice quality contrastively, and/or from languages in which the paralinguistic use of laryngeal voice quality differs from the unconstricted setting typically employed in Canadian English. Currently, no studies exist in the literature with which to compare the results of the present study.

### *6.1.2 Utterance Type*

Studies of infant speech production consistently note the development of syllabic utterances in the first year of life for children from a wide range of language backgrounds (Boysson-Bardies, 1999; Grégoire, 1937; Nathani et al., 1996; Roug et al., 1989; Stark et al., 1993), since babbling is considered to be a universal stage in speech and language development (MacNeilage, 1998; MacNeilage & David, 2001; MacNeilage et al., 2000; Oller 1978, 1980, 2000; Jakobson 1941/1968). As noted by Koopmans van Beinum and van der Stelt (1986), syllabic utterances are produced by the seventh month of life in 90% of infants. The failure to produce syllabic utterances by this time tends to be predictive of speech and language disorders (Oller et al., 1998, 1999). Based on information available at the time of writing, the Bai and English infants in the present study showed no signs of speech or language disorders. All infants in the study had produced some syllabic utterances by the seventh month of life, as shown in Figures 6-3 and 6-4 below, which illustrate the development of non-syllabic, mixed, and syllabic utterances for Bai and English infants, respectively.

**Figure 6-3. Utterance types of Bai infants, by age and percentage.**

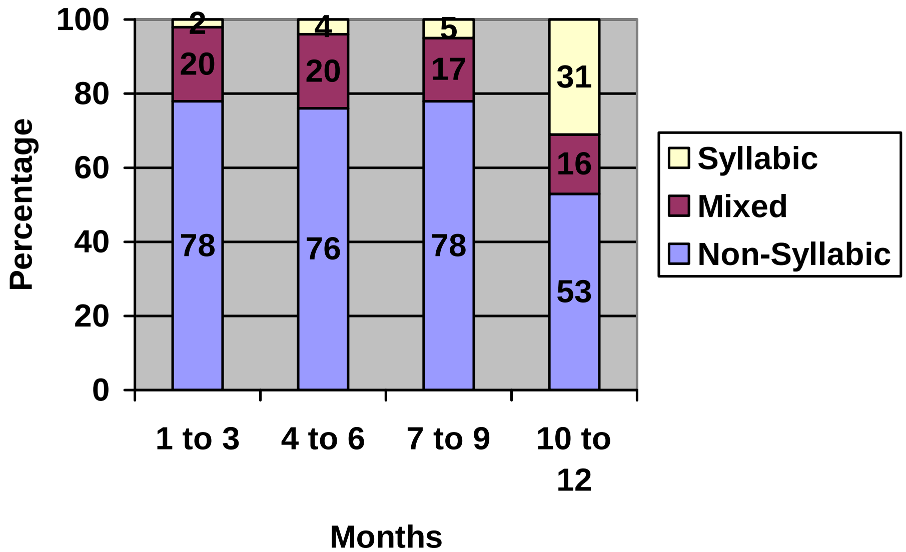
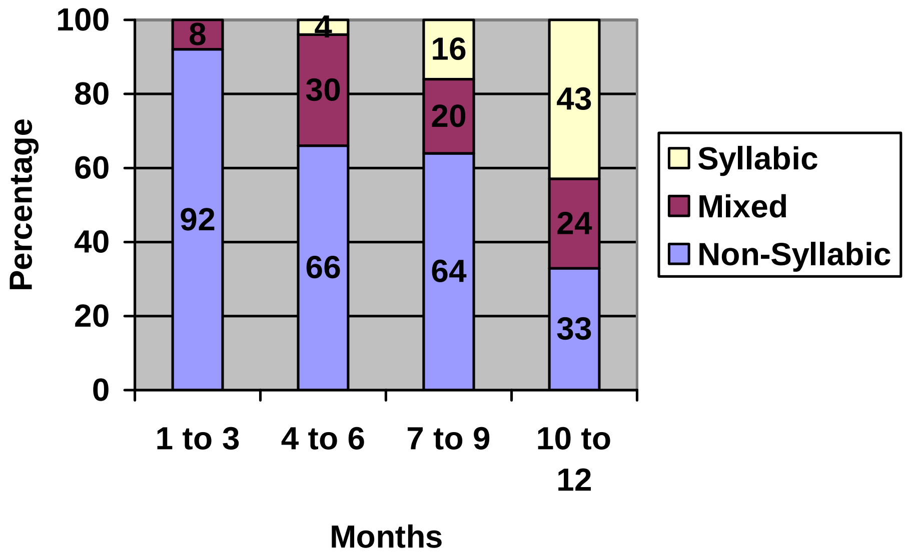


**Figure 6-4. Utterance types of English infants, by age and percentage.**




While the patterns found for the development of syllabic utterances confirm a universal tendency reported widely in the infant speech literature, there are striking differences in the rate of development of syllabic utterances between the two language groups in the second half of the year. Syllabic utterances increase significantly in months 7-9 among the English infants. By months 10-12, syllabic utterances make up 43% of English infants' total production. Among the Bai infants, the first syllabic utterances appear in small numbers early in the year, towards the end of months 1-3. However, these utterances do not show any significant increase until months 10-12, at which point they comprise 31% of infants' total production. Moreover, more than half of the Bai infants' total utterances in months 10-12 are non-syllabic, compared to only 33% for the English infants. This general pattern of development was found for each individual infant included in the Bai and English samples. Since all the infants in the study were developing typically, the explanation for the difference is unlikely to reflect a delay in speech and language development among the Bai infants. Rather, it is more likely that the difference in the rate of development of syllabic utterances stems from the influence of the ambient language.

Studies of the acquisition of tone consistently demonstrate that infants do not acquire specific lexical tones until the second year of life (Clumeck, 1977, 1980; Li & Thompson; Ota, 2003). Moreover, analyses of speech errors produced by young children, starting in the second year of life, show that infants who are learning a tone language make fewer errors in tone production than in the production of initial and final consonants (Zhu & Dodd, 2000). These patterns suggest that the information conveyed in tone is highly salient for infants who are learning tone languages, and that it may be at least as salient as the information encoded in the CV(C) structures that provide the segmental foundation for syllabic utterances and that are the primary carriers of contrastive information in non-tone languages. Given that by the age of six months, infants attend preferentially to lexical words (Shi & Werker, 2001), it seems likely that infants who are learning tone languages attend early on to pitch differences that create lexical contrast, including, in the case of the Bai infants, pitch-dependent changes in laryngeal voice quality. The salience of such variations within the language, combined with the fact that infants produce similar variations as a normal part of their phonetic development, may account for the greater prevalence of non-syllabic utterances in the Bai infants' production for most of the first year of life. While syllabic production is an essential part of speech development for infants from both language groups, non-syllabic utterances are not necessarily less "developed" or less "speech-like" than syllabic utterances for infants who are becoming attuned to pitch-dependent voice quality contrasts that can be explored in purely vocalic utterances, and later integrated into syllabic utterances. Thus, while infants from both language groups are naturally drawn to the production of syllabic utterances, the Bai infants may be just as strongly drawn to the production of non-syllabic utterances as an integral part of their phonetic and linguistic development, accounting for the greater prevalence of such utterances throughout the first year. By contrast, English infants, while engaging in considerable “vocal play” throughout the first year, may catch on fairly soon that the segmental information conveyed in syllables is more relevant to lexical contrast than pitch-dependent voice quality variations. Thus, early in development, syllabic vocalizations may exert a stronger pull for the English infants than for the Bai infants.

The distribution of syllabic, mixed, and non-syllabic utterances between the two languages also raises interesting questions about the integration of the laryngeal and oral vocal tracts in the course of speech and language development. Esling (2005) characterizes the laryngeal and oral vocal tracts as distinct domains of control. In the course of speech and language development, infants develop control over the laryngeal vocal tract before they develop control over the oral vocal tract. The integration of the functions performed by the laryngeal and oral vocal tracts is an essential part of speech and language development. The laryngeal vocal tract carries a heavy functional load in all languages. Arguably, however, it carries a heavier functional load in pitch-dependent tone languages, and an even heavier load in a language like Bai, which employs pitch-dependent laryngeal voice quality contrasts. This factor, in combination with the perceptual salience of laryngeal voice quality contrasts in Bai compared to English, may also serve to explain the slower development of syllabic utterances among the Bai infants. To produce syllabic utterances that begin to resemble the syllabic utterances in their ambient language, Bai infants must develop the capacity to produce rapid alternations in laryngeal voice quality from syllable to syllable. This capacity requires a higher degree of laryngeal control than is required for a language in which syllabic utterances are normally produced in a single, unconstricted voice quality. Thus, Bai infants may be motivated to invest more time than English infants in developing refined laryngeal control before they concentrate on the production of syllabic utterances, which represent a comparatively advanced integration of the oral and laryngeal vocal tracts.

While the above discussion is a plausible response to the findings of the present study within the theoretical perspective adopted herein, it is also quite speculative. The explanations advanced above for the development of syllabic and non-syllabic utterances in the Bai and English infants need to be explored in larger studies of Bai and English infants, including follow-up studies on the development of laryngeal voice quality in the second year of life among Bai and English infants. Tracking the emergence of pitch-dependent voice quality variations at the syllabic level in the second year of life, the period when tones are acquired, may make the phonetic developmental processes of the first year more transparent, particularly for the Bai infants. Currently, there are no systematic studies of the development of laryngeal voice quality in tone or non-tone languages that could shed light on the findings reported herein.

### *6.1.3 Laryngeal Voice Quality in Mixed and Syllabic Utterances*

In his work on infant speech production, Oller (1978, 1980, 2000) has consistently commented on the importance of modal voice in speech development generally, and in the development of syllabic utterances in particular. Given Oller's strong influence on studies of infant speech production, it is worth quoting at some length his account of the relationship between the development of modal voice and canonical syllable production:

> Infraphonological achievements are designated in terms of principles or abstract characteristics of canonical syllables … . Quasivowels of the first stage manifest only one of these properties, normal phonation. In gooing, during the second stage, normal phonation is combined with another property, articulation, or movement of the vocal tract. In gooing, vocal tract movement is usually limited in extent, which is to say that the mouth is not moved very far from its rest position. Nonetheless the limited articulation of gooing demonstrates that infants in this stage are able to produce normal phonation and articulation simultaneously. In the third stage, the movements of the tract are made more extensive as the infant explores the range of possible articulatory postures, producing wide swings of vocal tract movement and concomitant full resonance patterns. In marginal babbling of the third stage, the infant incorporates the canonical features of normal phonation, full resonance, and fully articulated movements between full vowels and consonant-like margins in primitive syllables that are only one step away from being fully well-formed. In canonical babbling, during the fourth stage, the last feature is added to the package, as the infant produces fully articulated sequences with rapid transitions between closures and openings, yielding well-formed syllables. (Oller, 2000: 67)

In terms of production, Oller stands out among infant speech researchers for his contribution to our understanding of speech and language development as a continuous process that begins at birth. However, his characterization of the role of modal phonation in speech and language development may be oversimplified, and in some respects, inaccurate. Based on the findings of the present study, it is true that infants' speech and language development correlates with an increase in the frequency of production of modal voice (modal voice is the most frequently produced of the unconstricted voice quality settings, comprising 81% of Bai infants' unconstricted vocalizations, and 76% of English infants' unconstricted utterances). However, it is not true that the production of modal voice is necessarily associated with the emergence of oral consonants or with the emergence of syllabic utterances. Rather, based on the findings of the present study, it seems that oral consonants and syllabic utterances first emerge in utterances produced with constricted and dynamic laryngeal voice quality settings, and that practice in these settings produces the type of laryngeal control that allows for the greater production of modal voice and the increased production of syllabic utterances. Thus, increased production of modal voice is the *effect* of developing laryngeal control, which originates in the production of constricted and dynamic utterances, not its *cause*. In turn, increased laryngeal control promotes an increase in the frequency of syllabic utterances, regardless of laryngeal voice quality, though it may be easier to produce a rapid sequence of canonical syllables in modal voice than in constricted and dynamic settings, with the latter being the most difficult to achieve.

A review of the findings on the distribution of laryngeal voice quality settings in mixed and syllabic utterances may help to clarify the points raised above. Oller's description of the degree of oral vocal tract closure in "gooing" essentially corresponds to the degree of oral vocal tract closure featured in the category of "mixed" utterances employed in this study. Figures 6-5 and 6-6 below show the distribution of laryngeal voice quality in mixed utterances for Bai and English infants, respectively. In both languages, mixed utterances begin to appear towards the end of months 1-3. When they first appear in months 1-3, they are most often produced in constricted utterances for infants from both language groups.  For the remainder of the year, mixed utterances are most likely to be produced in vocalizations with dynamic laryngeal voice quality. They are least likely to be produced with unconstricted voice quality settings, including modal voice, though infants from both language groups begin to produce mixed utterances with unconstricted voice quality settings in significant numbers in months 7-9. Thus, to use

Oller's term, "gooing" in modal voice is a comparatively late development, compared to "gooing" in constricted and dynamic laryngeal voice quality settings. This pattern may be explained by the notion of "laryngeal priming," introduced by Esling et al. (2006) to explain the possible role of laryngeal constriction—the earliest type of utterance controlled by infants—in priming the development of the oral vocal tract, and in integrating the strictures produced in the laryngeal and oral vocal tracts over the course of speech and language development.

**Figure 6-5. Laryngeal quality in Bai mixed utterances, by percentage.**

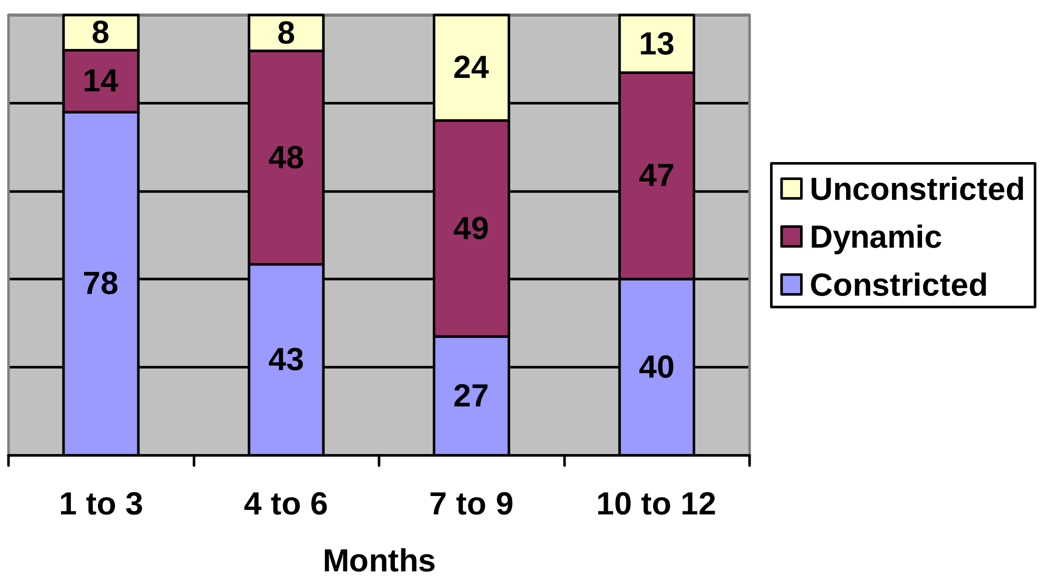
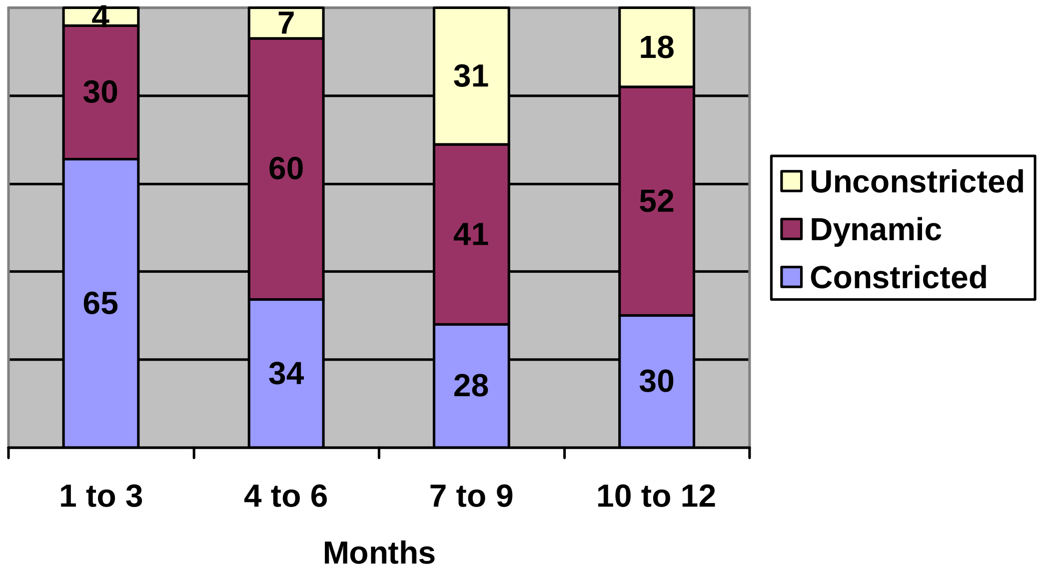


**Figure 6-6. Laryngeal quality in English mixed utterances, by percentage.**




The patterns seen in the distribution of laryngeal voice quality in syllabic utterances place Oller's evaluation of the importance of modal voice into further question. Figures 6-7 and 6-8 illustrate the distribution of laryngeal voice quality settings in syllabic utterances for Bai and English infants, respectively. At first glance, results for the Bai infants might seem to support Oller's (2000) theoretical perspective, given that the earliest syllabic utterances produced by these infants are often produced with unconstricted laryngeal voice quality settings. However, it is important to recall that the number of syllabic utterances produced by Bai infants is very low in months 1-3, 4-6, and 7-9, rising significantly only in months 10-12. The distribution of laryngeal voice quality in the babbling of the English infants, shown in Figure 6-8, may be more telling, as these infants produce increasing numbers of syllabic utterances throughout the course of the first year, starting in months 4-6.

**Figure 6-7. Laryngeal quality in Bai syllabic utterances, by percentage.**

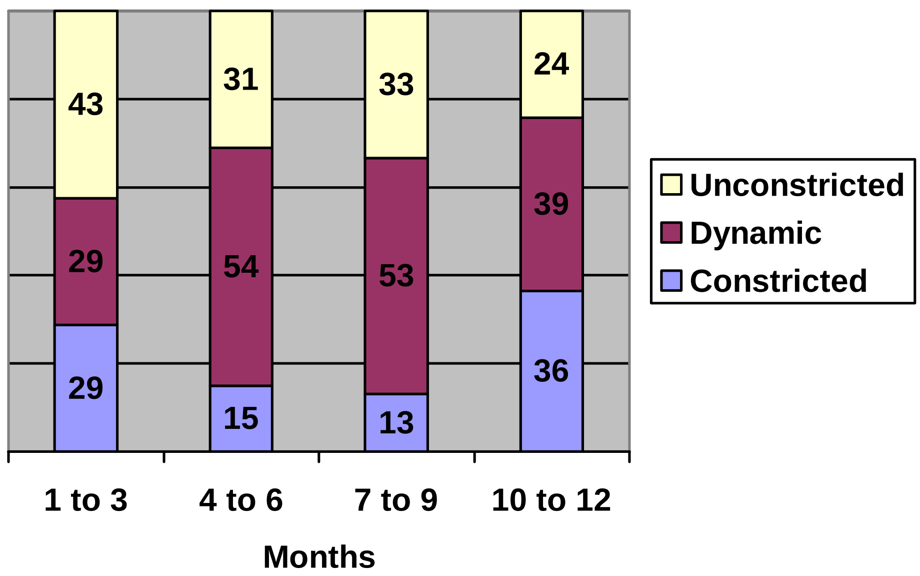
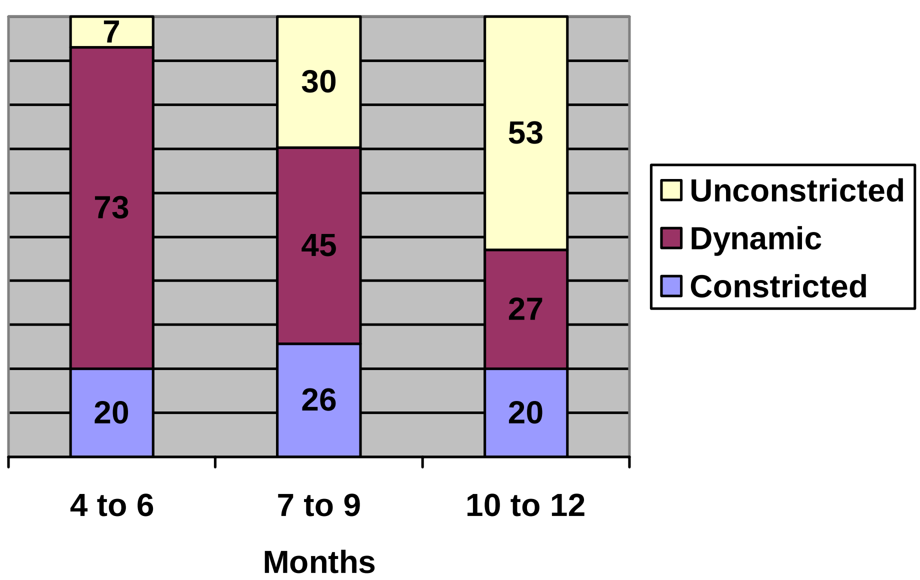



**Figure 6-8. Laryngeal quality in English syllabic utterances, by percentage.**




As shown above, when syllabic utterances first begin to appear in the utterances of the English infants, they are most commonly produced in constricted or dynamic laryngeal voice quality settings. In months 4-6, most syllabic utterances produced by the

English infants reflect the timing characteristics of Oller's definition of "marginal babbling," that is, babbling in which the transitions from the consonants to the vowels are slow. It is likely that syllabic utterances from this period are "marginal," in terms of their timing characteristics, partly *because* they are produced in constricted and dynamic laryngeal voice quality settings, and that the impetus to produce a rapidly articulated series of syllables that reflects the timing characteristics of syllabic productions in the infants' ambient language *stimulates* the English infants to produce more modal voice. In other words, English infants do not "master" modal voice prior to developing syllabic utterances, as suggested by Oller. Rather, following the emergence of oral vocal tract control in constricted and dynamic mixed utterances, English infants begin to produce constricted and dynamic syllabic utterances, and then go on to develop increased control of those settings over the course of the year, along *with* unconstricted settings such as modal voice. It may require a lesser degree of laryngeal control to produce canonically timed syllables in unconsticted settings than in constricted—and especially dynamic—laryngeal voice quality settings. An examination of mean length of utterance of syllabic utterances by laryngeal voice quality is illustrative.

Table 6-1 and Figure 6-9 below present the mean length of syllabic utterances by laryngeal voice quality for both languages. Table 6-2 and Figure 6-10 below show the mean length of syllabic utterances for English utterances by laryngeal quality for months 4-6, 7-9, and 10-12 (the breakdown by age is not provided for Bai infants, because the numbers for earlier age periods are too small within each category to be informative). As shown in Figure 6-9, for both language groups, the mean length of constricted and unconstricted syllabic utterances is approximately the same, while the mean length of dynamic utterances is longer, most likely because of the inherent physical challenge of alternating between constricted and unconstricted settings. Moreover, as demonstrated in Figure 6-10 for the English infants, the earlier in phonetic development the syllabic utterance, the longer the mean length across all laryngeal voice quality categories. For syllabic utterances, shorter mean length corresponds to increased laryngeal control, since it tends to coincide with a more rapid articulation of syllables.

**Table 6-1. Mean length of syllabic utterances, by language.**

| Laryngeal Voice Quality | Bai (s.d.) | English (s.d.) |
|---|---|---|
| Constricted | 1.56 (1.22) | 2.21 (2.21) |
| Dynamic | 3.45 (3.52) | 4.28 (3.42) |
| Unconstricted | 1.57 (1.17) | 2.22 (1.87) |

**Figure 6-9. Mean length of syllabic utterances, by language.**

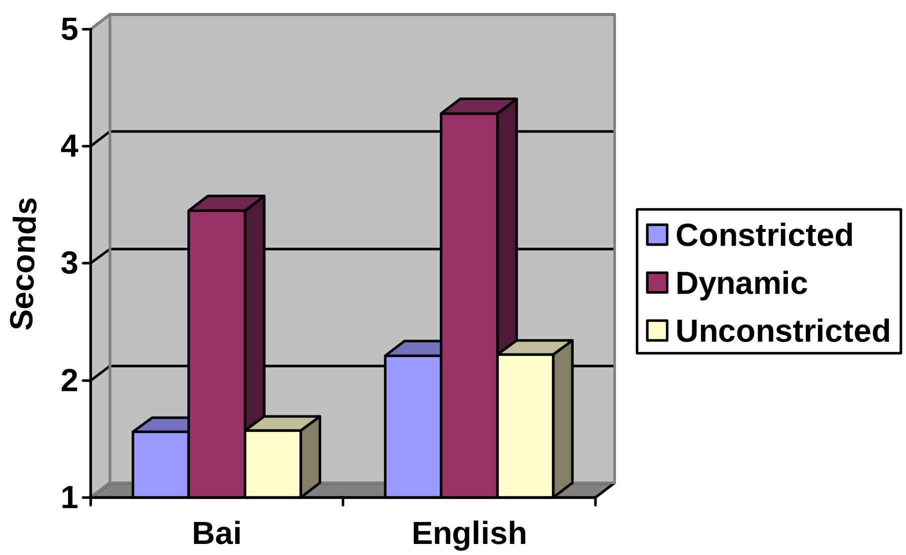
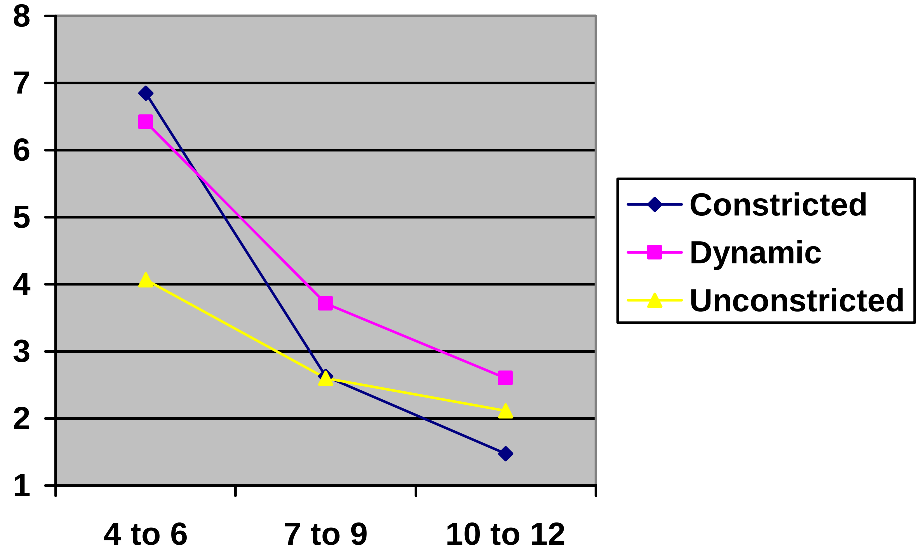



**Table 6-2. Mean length of English syllabic utterances, by laryngeal quality and age.**

| Age | Constricted | Dynamic | Unconstricted |
|---|---|---|---|
| Months 4-6 | 6.86 (5.35) | 6.42 (2.49) | 4.07 (0) |
| Months 7-9 | 2.63 (1.45) | 3.71 (2.84) | 2.60 (2.15) |
| Months 10-12 | 1.48 (1.16) | 3.89 (3.75) | 2.11 (1.81) |

**Figure 6-10. Mean length of English syllabic utterances, by laryngeal quality and age.**




The capacity to produce syllabic utterances in dynamic and constricted settings may develop earlier than it does in unconstricted settings, but it may take a greater degree of laryngeal control to produce canonically timed sequences in these settings. When the English infants begin to produce longer babbling sequences, in months 4-6, they do not yet have the degree of laryngeal *or* oral control to produce canonically timed syllables, hence the relatively long mean length of utterance for syllabic vocalizations at this time across all laryngeal voice quality settings. With increased practice of laryngeal control in both syllabic and non-syllabic utterances, however, the English infants begin to produce shorter syllabic utterances across all laryngeal voice quality categories. While dynamic syllabic utterances remain the longest, on average, of all syllabic utterances, the gap between the mean length of constricted, dynamic, and unconstricted syllabic utterances narrows with time. By the time the Bai infants begin to produce syllabic utterances in significant numbers in months 10-12, they have already developed a degree of laryngeal control that allows them to produce syllabic utterances with approximately the same timing characteristics as the syllabic utterances produced by the English infants across all laryngeal voice quality categories, despite the fact that they have not concentrated on producing syllabic utterances in the earlier months to the same extent as the English infants (none of the Bai syllabic utterances are as long as those observed in the early syllabic utterances of the English infants). Among Bai infants, the mean length of constricted, dynamic, and unconstricted utterances is 1.79 sec, 3.91 sec, and 1.53 sec in months 10-12, respectively, which is roughly comparable to the corresponding values observed for the English infants during this same period, at 1.83 sec, 3.89 sec, and 2.11. Notably, however, the Bai infants have developed this degree of laryngeal control almost exclusively through the medium of mixed and non-syllabic utterances, most often produced with constricted and dynamic laryngeal voice quality settings. This strategy allows them to simultaneously practice skills in laryngeal voice quality that may be important in their acquisition of the tone system in Bai.

Thus, while Oller is correct in observing that laryngeal control is important to the production of syllabic utterances, he may be wrong in assuming that laryngeal control is primarily expressed through the production of modal voice. By the time that Bai and English infants produce syllabic utterances in significant numbers in months 10-12, they both have considerable control of the laryngeal and oral vocal tracts, though they may have gained this control in slightly different ways, according to the salience of different features in their respective target languages. Also, by this time, infants from both language groups are producing syllabic utterances with laryngeal voice quality settings that are beginning to resemble the use of such settings in their target languages. The syllabic utterances of Bai infants are mostly constricted and dynamic, while the syllabic utterances of English infants are mostly unconstricted. These differences likely reflect the infants' attunement to the use of laryngeal voice quality in Bai and English. However, to confirm these patterns, further research is necessary on more infants who are learning Bai and English, and on more infants who are learning languages that make contrastive use of laryngeal voice quality.

Currently, there is no basis for comparison of the results of this study with other research on infant speech production in the literature. To date, no other studies of laryngeal voice quality in syllabic utterances have focused on infants who are learning a register tone language. Furthermore, most research on English syllabic utterances excludes babbling produced with laryngeal constriction, because these utterances are difficult to analyze instrumentally (see, for example, Boysson-Bardies et al., 1984 and Hallé & Vihman, 1991). However, in considering the effect of the ambient language on the production patterns documented in this study, it may be possible to make reasonable speculations based on the perceptions of laryngeal voice quality in infant vocalizations by Bai and English adults. To that end, we now turn to a brief discussion of the findings of the perceptual study reported in Chapter 5.

## 6.2 Perception

In this section, I discuss the findings of the perception component of the present study. As with the discussion of the production results, the discussion of adult perception of infant vocalizations addresses the five questions that motivated this part of the study. These questions, repeated for convenience from Chapter 1, are:

- **<u>Question 1</u>:** Can phonetically untrained adults systematically evaluate infant vocalizations outside of specific relationships with infants they know?
- **<u>Question 2</u>:** Are phonetically untrained adults sensitive to laryngeal voice quality features in evaluating infant vocalizations?
- **<u>Question 3</u>:** Are there universal patterns in phonetically untrained adults' perceptions of laryngeal voice quality features in infant vocalizations?
- **<u>Question 4</u>:** What is the relationship between adults' ratings of infant vocalizations and universal processes of infant phonetic development in production? Do higher ratings correlate with features that infants control later in the process of phonetic development?
- **<u>Question 5</u>:** Do language-specific factors affect phonetically untrained adults' perceptions of laryngeal constriction in infant vocalizations? If language-specific factors exist, do these factors relate systematically to language-specific patterns in the use of laryngeal constriction in infants' production?

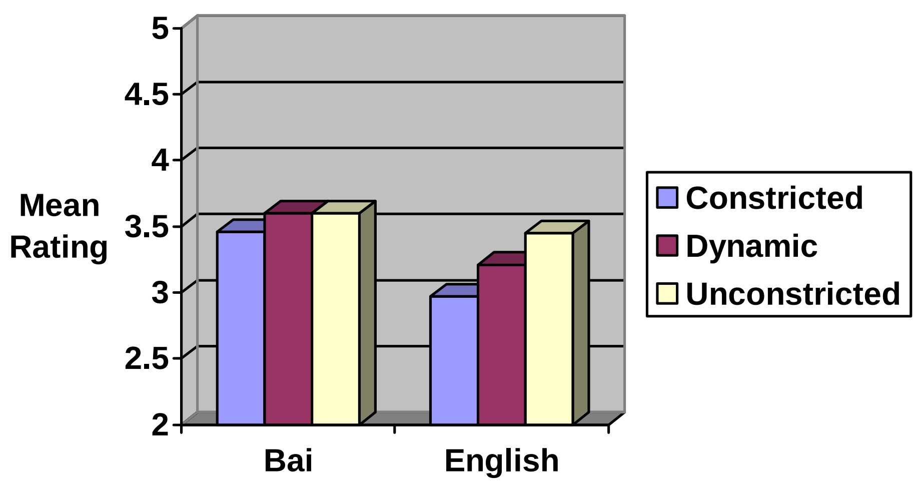
Sections 6.2.1 and 6.2.2 discusses universal and language-specific patterns in adults' perceptions of laryngeal voice quality and utterance type, respectively. Section 6.2.3 discusses the relationship between laryngeal voice quality and utterance type in the perceptions of Bai and English adults. As in the discussion of the results of the production study, figures that were presented in Chapter 5 are reproduced in this chapter as relevant, for convenience. The question of how adults' perceptions relate to universal and language-specific patterns in Bai and English infants' production, as articulated in research question 5 above, will be integrated throughout the discussion.

### *6.2.1 Laryngeal Voice Quality*

To my knowledge, the perceptual study reported in this dissertation is the first study of adults' perceptions of laryngeal voice quality in syllabic and non-syllabic non-distress vocalizations of infants. It is also the first study that compares the perception of laryngeal voice quality between two groups whose languages differ significantly in their use of laryngeal voice quality settings. Based on existing knowledge, while all languages use laryngeal voice quality variations for paralinguistic purposes (Esling, 1994, 2000; Laver, 1980, 1994), few languages of the world use laryngeal voice quality settings contrastively at the syllabic level (Esling, 2000; Esling & Edmondson, 2002). This exploratory study compares adult perceptions of laryngeal voice quality by adults who speak English, a language that uses laryngeal voice quality variations for paralinguistic purposes, and by adults who speak Bai, a language that employs laryngeal voice quality variations contrastively in its register tone system. To date, studies that focus on adults' perception of laryngeal voice quality in infant vocalizations are restricted to speakers of Indo-European languages, and to perceptions of infant crying (Bisping et al., 1990; Frodi, 1985; Frodi & Lamb, 1980; Frodi & Senchak, 1990; Murray, 1985; Zeskind, 1987; Zeskind & Shingler, 1991). While these studies often highlight adults' negative reactions to laryngeal constriction, it is not clear that laryngeal constriction in non-distress vocalizations would evoke similar responses from adults, particularly from adults whose native language employs laryngeal constriction to create lexical contrast.

Figure 6-11 below illustrates the overall pattern in Bai and English adults' ratings of laryngeal voice quality. Generally speaking, these ratings reflect the distribution of laryngeal constriction in the two languages. Given that Bai lexical items may include syllables that include laryngeal constriction, that alternate between constricted and unconstricted settings, and that are unconstricted throughout, it is reasonable that Bai listeners might not express a strong preference for any of the utterances in this study, purely on the basis of laryngeal voice quality. By contrast, among Canadian English speakers, unconstricted voice quality settings are the norm. Constricted voice quality settings tend to be associated with the expression of strong—often negative—emotion (Esling, 1994; Laver, 1980), and/or with voice pathologies (Hollien, 1974). Thus, it is not surprising that English listeners should express an overall preference for unconstricted utterances, and that they should least prefer constricted utterances.

**Figure 6-11. Mean ratings of laryngeal quality, by language.**

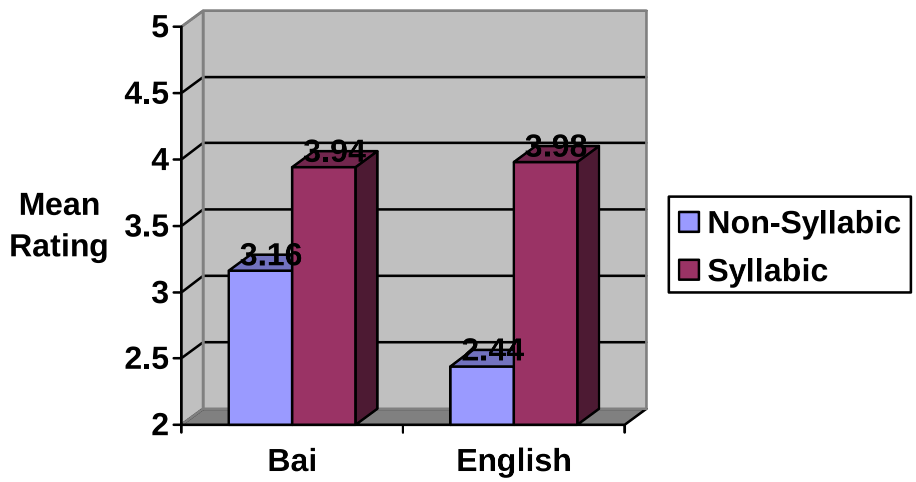



Generally speaking, adult ratings from the two languages relate systematically to the overall distribution and development of laryngeal constriction in infant utterances throughout the first year, particularly for English. As reviewed in Section 6.1.1 above, and throughout Chapter 4, while the incidence of laryngeal constriction decreases throughout the year, the decline is more marked among the English infants than among the Bai infants. Conversely, the incidence of unconstricted utterances increases for infants from both language groups, but increases to a greater degree among the English infants. While the overall development of laryngeal constriction in infants' utterances is strongly influenced by physiological constraints, particularly in the first half of the year, the pattern observed in the second half of the year may reflect the influence of the ambient language. English infants' production of constricted and unconstricted utterances may be a response to English adults' aversion to constricted sounds, and to their attraction to unconstricted sounds. It is notable that in the qualitative component of the perceptual study, several English adults reported actively discouraging the proliferation of constricted sounds, and encouraging the production of unconstricted sounds in their relationships with infants, a behavioral pattern that is consistent with Locke's (2006) hypothesis that parents and other caregivers select for vocal behavior in child-rearing.

#### *6.2.2 Utterance Type*

The existing literature on adult perceptions of infant vocalizations suggests that adults strongly prefer syllabic over non-syllabic utterances (see, for example, Bloom & Lo, 1990; Bloom et al., 1993; Goldman, 2001) and that phonetically untrained adults from a wide range of educational and socio-economic backgrounds recognize the onset of babbling in infants, even if they lack the phonetic vocabulary to describe it (Oller et al., 2001). As illustrated in Figure 6-12 below, the results of this perceptual study are consistent with existing research in this respect: Bai and English adult listeners clearly award higher ratings to syllabic utterances than to non-syllabic utterances, reflecting the importance of syllabic utterances in all human languages (Oller, 2000) and possibly, a preference for reduplicative patterns among primates generally (Elowson et al., 1998a, 1998b).

**Figure 6-12. Mean ratings, by utterance type and language.**




It is, however, noteworthy that while Bai and English adult mean ratings for syllabic utterances are approximately the same, Bai adults rate non-syllabic utterances more highly, on average, than do English adults. Moreover, adults' preferences are systematically related to differences in the production of syllabic versus non-syllabic utterances between Bai and English infants, as reported in Section 6.1.2 above, and throughout Chapter 4. As discussed, Bai infants do not produce syllabic utterances in significant numbers until months 10-12, despite the fact that they appear at least as early in the first year for these infants as they do for English infants. My proposed explanation for this pattern is that laryngeal voice quality variations are more salient in the ambient language for Bai-learning infants than for English-learning infants, and that, possibly, Bai infants may produce more vocalic utterances than English listeners in response to this salience. It is possible that while Bai and English adults both favour syllabic utterances, Bai adults also intuitively recognize the role of non-syllabic utterances in infants' development of the laryngeal control necessary to acquire the tones of the language—hence their comparatively higher ratings for non-syllabic utterances, compared to English adults.

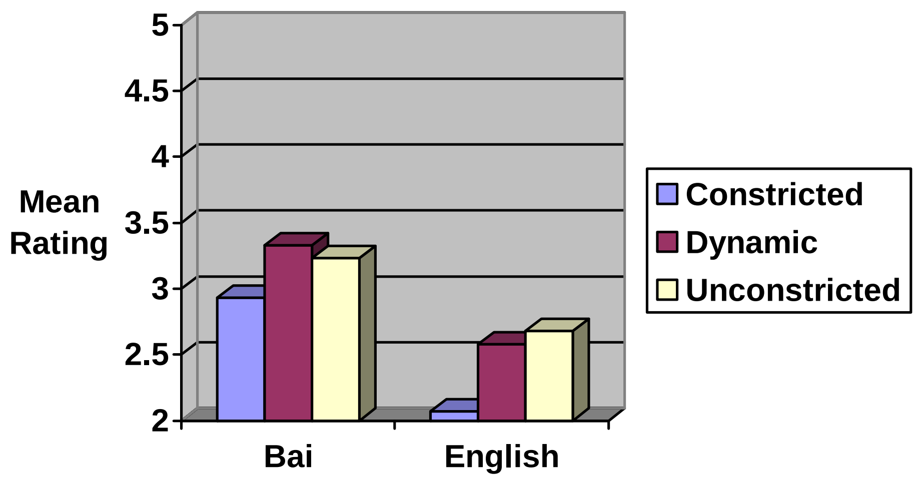
By contrast, English adults may generally perceive non-syllabic utterances as immature expressions of emotion that can be indulged for a time in infancy, but that are best transferred to the syllabic and, later, the lexical domains as soon as possible in development, or as "mere" play, the value of which is often underestimated in many learning processes, particularly in cultures that are strongly influenced by Northern European and/or Protestant values. In the qualitative component of the study, many English listeners expressed ambiguous attitudes towards constricted and dynamic non-syllabic utterances that they perceived to reflect emotional expression, playing, or "letting it all hang out." English infants' relatively greater production of syllabic utterances may reflect the influence of these attitudes, in combination with the greater salience of syllabic structures in carrying lexical meaning in their ambient language.

### 6.2.3 Laryngeal Voice Quality by Utterance Type

#### *6.2.3.1 Non-Syllabic Utterances*

Figure 6-13 illustrates Bai and English ratings of laryngeal voice quality in non-syllabic utterances. This figure illustrates the relatively higher ratings awarded by Bai listeners for non-syllabic utterances, as discussed above. The figure also shows differences in relative preferences for different laryngeal voice quality settings in non-syllabic utterances, along with some common trends.

**Figure 6-13. Ratings of non-syllabic utterances, by laryngeal quality and language.**




Listeners from both language groups give the lowest ratings to constricted non-syllabic utterances. These utterances are produced in large numbers by Bai and English infants throughout the year, but constitute the vast majority of utterances produced at the beginning of life, in months 1-3. Adult ratings likely reflect an awareness that these utterances are produced by infants early in life, before they show any obvious signs of language development. However, the especially low ratings awarded by the English listeners may reflect a cultural aversion to laryngeal constriction, while the comparatively high ratings assigned to these utterances by the Bai listeners may reflect the importance of laryngeal constriction in the target language. In the qualitative component of the study, many Bai participants specifically cited the importance of constricted sounds in Bai. However, many participants also suggested that such sounds were “less advanced” or expressive of the infants’ desire for attention. Bai listeners’ ratings of these sounds ranged widely, probably in response to these conflicting factors in the evaluation of laryngeal constriction, particularly when it is presented to them outside of a particular communicative context.

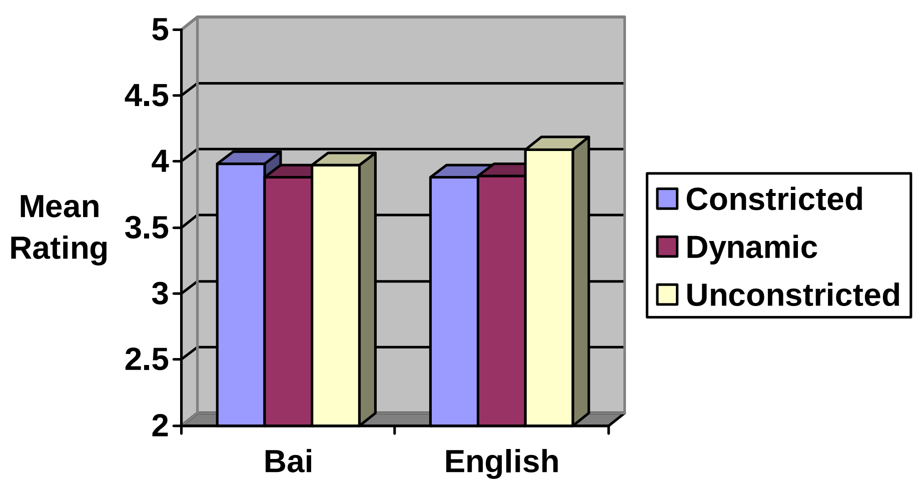
Ratings for dynamic and unconstricted non-syllabic utterances were notably higher than ratings for constricted non-syllabic vocalizations for listeners from both language groups. As documented in Section 6.1.1 above, and throughout Chapter 4, these utterances appear in significant numbers later in infants’ development than do constricted non-syllabic utterances. Adults’ relatively higher ratings for such sounds may reflect an intuitive awareness of their importance in infants’ phonetic development. Among non-syllabic utterances, Bai adults rated dynamic utterances the highest, on average, while English adults rated unconstricted utterances the highest. These slight differences may reflect the importance of laryngeal voice quality variations in Bai, on the one hand, and the English cultural preference for unconstricted utterances, on the other.

Many factors relating to pitch range, pitch contours, and the timing of alternations in pitch and/or voice quality affected the ratings of individual items across, between, and within the two language groups. The production component of the present study does not explore these factors in detail. Thus, it is not possible to say whether adults’ preferences in these respects relate systematically to patterns found in infants’ utterances, over time, or within language groups. However, the identification of these factors may assist in more detailed studies of the development of Bai and English infants’ utterances, both in the first and second years of life. Based on the level of detail considered in the present study of infants’ production, the general patterns in the adults’ ratings of non-syllabic utterances relate systematically to infants' production patterns primarily in terms of stages of development, mediated by cultural and linguistic preferences. Non-syllabic utterances that appear early in life are rated lower than non-syllabic utterances that develop later by adults from both language groups.

*6.2.3.2 Syllabic Utterances*

Ratings for syllabic utterances were broadly similar for Bai and English adults, as illustrated in Figure 6-14 below. Listeners from both language groups gave the highest ratings to syllabic utterances in which infants produced a series of rapidly articulated syllables in a short period of time—an ability cultivated by most Bai and English infants only in months 10-12, whether they started to produce babbling sequences early in the year, or later.

**Figure 6-14. Ratings of syllabic utterances, by laryngeal quality and language.**



As with the ratings for non-syllabic utterances, many additional factors affected the ratings of the Bai and English adults, across, between, and within languages. Listeners from both language groups gave lower ratings to utterances that were relatively monotone in pitch, with Bai men and English women being especially averse to such sounds. In their ratings of syllabic and non-syllabic sounds, Bai men prized the ability to produce alternations in pitch and/or voice quality, regardless of whether such alternations occurred in constricted, dynamic, or unconstricted settings, or in syllabic or non-syllabic utterances. While Bai women also favoured such alternations, their preferences systematically differed according to whether the utterance was syllabic or non-syllabic. In non-syllabic utterances, Bai females preferred slower transitions in pitch and/or voice quality. By contrast, in syllabic utterances, Bai women gave higher ratings to utterances that alternated rapidly in pitch and/or voice quality from syllable to syllable, as might occur in the Bai language. English listeners had a general preference for unconstricted syllabic utterances produced within a moderate pitch range, in a single voice quality.

In some general respects, the preferences of the Bai and English adult listeners roughly correspond to the patterns observed in the production of syllabic utterances for infants from the two language groups. For example, in months 10-12, the majority of syllabic vocalizations produced by English infants are unconstricted, possibly reflecting the preference for unconstricted syllabic utterances in the ambient language. By contrast, the syllabic utterances produced by the Bai infants are relatively evenly distributed among constricted, dynamic, and unconstricted laryngeal voice quality categories, reflecting the adult Bai listeners' relative lack of preference for any given sub-category of syllabic utterance, on the basis of laryngeal voice quality alone. While Bai adults' ratings for syllabic dynamic utterances are lower, on average, than for syllabic constricted and syllabic unconstricted utterances, these ratings appear to be affected by the timing characteristics of the dynamic syllabic vocalizations included in the sample. These timing features may, in turn, reflect the inherent challenges infants face in acquiring the ability to produce a series of rapidly articulated syllables that alternate between constricted and unconstricted laryngeal voice quality features, compared to producing such a series within constricted or unconstricted syllabic utterances.

The more specific perceptual preferences of the Bai and English listeners regarding the distribution of pitch and/or voice quality in syllabic utterances could help to guide more fine-grained phonetic studies of Bai and English infants' production in the first and second years of life. These preferences could also serve to guide the selection of infant speech stimuli in larger, more controlled studies that test the perceptions of adult listeners from the two language groups.

It is important to note that while there was considerable variation in the ratings of individual items in every sub-category of utterance in the perceptual test, the variability was systematic across, between, and within languages, in nearly all cases. It is also important to note that in rating the range of infant utterances included in this perceptual test, Bai and English adults did not specifically prefer utterances produced by infants learning their own language. Rather, Bai and English adults appeared to rate infants' utterances on the basis of a set of desirable features, most of which were shared across the two languages, and some of which were particular to patterns in their respective languages and/or cultures. Adults' ratings most strongly reflected their intuitive knowledge of stages in infant phonetic development, as seen in the overall preference for rapidly articulated syllabic utterances, which develop comparatively late among infants from both language groups; and the relative aversion to constricted non-syllabic utterances, which are the earliest vocalizations produced by infants. Some language-specific preferences do systematically relate to differences in production patterns between infants from the two language groups. For example, the general English preference for unconstricted sounds is reflected in the higher incidence of such sounds in syllabic and non-syllabic utterances produced by English infants, compared to the somewhat lower incidence of such sounds in the utterances produced by Bai infants. Whether language-specific preferences with respect to finer-grained phonetic details, such as pitch range, pitch contours, and the timing of pitch and voice quality variations are reflected in Bai and English infants' production in the first year will require further study. However, based on existing studies of prosodic development (see Snow & Balog, 2002, for a review), which do not address laryngeal voice quality, it seems likely that Bai and English infants' productions would not begin to strongly reflect these finer-grained differences until the second year of life.

# Chapter 7

# CONCLUSION

In this chapter, I summarize the main results of the production and perception components of the present study, and describe potential directions for future research. In Section 7.1, I summarize the results of the production study and their implications for current understandings of phonetic development in the first year of life. In Section 7.2, I outline the main findings of the perceptual study. In Section 7.3, I describe how this exploratory research might inform future research into infant speech development and studies of adult perceptions of infant vocalizations.

## 7.1 Infant Speech Production

This dissertation is a report of the first phonetically grounded study of the development of laryngeal voice quality throughout the first year of life, building on the preliminary findings of the InSpA Team at the University of Victoria (Benner et al., 2007; Bettany, 2004; Esling et al., 2004a, 2004b, 2004c, 2006; Grenon et al., 2007). This study examines the development of laryngeal voice quality in the utterances of Bai and English infants. In comparing these two groups of infants, this study is also the first in the infant speech literature to compare the development of laryngeal voice quality among infants who are learning languages that differ significantly in their use of laryngeal voice quality. In English, as in the the majority of languages, laryngeal voice quality serves primarily as the carrier of paralinguistic and indexical information (Esling, 1994, 2000; Laver, 1980, 1994). However, a small number of languages, like Bai, a Tibeto-Burman language spoken in Yunnan, China, employ laryngeal voice quality contrastively as part of a register tone system (Esling & Edmondson, 2002). It is of interest to know whether there are universal patterns in the development of laryngeal voice quality across languages, what role laryngeal mechanisms play in speech and language development, and whether laryngeal voice quality develops differently in infants who are learning a language that employs laryngeal voice quality distinctively.

The production study reported in this dissertation is grounded in the theoretical perspective developed by Esling and his colleagues, which is based on extensive laryngoscopic observations of the larynx in the production of the full range of laryngeal sounds employed in languages of the world (Carlson et al., 2004; Edmondson et al., 2005; Edmondson & Esling, 2006; Esling, 1996, 1998, 1999a, 1999b). This theoretical perspective provides a unified account of the physiological mechanisms involved in laryngeal constriction, which were, until quite recently, poorly understood (Catford, 1977; Maddieson & Ladefoged, 1996). This body of work allows us to see the full range of laryngeal sounds produced by infants and adults alike in terms of degrees of constriction within the larynx. Esling’s (2005) model of the vocal tract, which characterizes the laryngeal vocal tract as an articulator distinct from the oral vocal tract, also allows us to conceptualize speech and language development as a gradual integration of two separately controlled articulators.

Many previous infant speech researchers have described the prevalence of laryngeal constriction in the early utterances of infants (see, for example, Stark et al., 1975 and Koopmans-van Beinum & van der Stelt, 1986), but they have described these sounds impressionistically (Keating, 1980; McCune et al., 1996; Oller, 1978, 1980, 2000; Roug et al., 1989), or have categorized them in diverse ways that obscure the underlying laryngeal mechanisms that unite such sounds as phonetic categories (Nathani et al., 2006). Other researchers have excluded these sounds from analysis, potentially concealing their potential role in the process of speech and language development (Hallé & Vihman, 1991; McCune & Vihman, 2001). This study includes and categorizes the full range of non-distress utterances produced by infants in the first year of life in terms of laryngeal voice quality categories, highlighting universal and language-specific features of development, as well as the relationship between laryngeal and oral control.

Based on the findings of this study, we may say that infants from different language backgrounds produce primarily constricted sounds in the first months of life and go on to produce an increasing proportion of unconstricted sounds throughout the remainder of the first year of life. In months 4-6, in concert with physiological developments that afford infants freedom to produce a wider range of sounds in the larynx, infants begin to produce dynamic sounds—vocalizations that alternate between constricted and unconstricted laryngeal voice quality settings—in increasing numbers. The production of dynamic sounds appears to play a significant role in infants' development of laryngeal control over the first year of life and occurs prior to their developing capacity to produce unconstricted sounds in significant numbers. Also, in the first months of life, infants produce their first oral sounds, usually in the context of utterances produced with constricted laryngeal voice quality. Over time, infants develop increasing control over the laryngeal and oral vocal tracts, allowing them to produce syllabic utterances featuring rapidly articulated canonical syllables towards the end of the first year, in months 10-12.

In terms of laryngeal voice quality, infants' utterances begin to reflect features of the ambient language in the first year of life. While all the infants in this study produced an increasing proportion of unconstricted utterances throughout the first year, among the English infants, unconstricted utterances increased more quickly and to a greater extent, possibly reflecting an attunement to the use of laryngeal voice quality in the ambient language. Similarly, while the incidence of constricted utterances declined throughout the year for infants from both language groups, these vocalizations did not decline as quickly or to the same degree among the Bai infants, whose ambient language features laryngeal constriction. Finally, among English infants, dynamic utterances increased significantly in months 4-6, during the period of "vocal play" (Oller, 2000), but then decreased slightly for the remainder of the year. In line with the pattern observed in the English infants, among Bai infants, the incidence of dynamic utterances increased in months 4-6, but then continued to slowly increase for the balance of the first year, possibly reflecting the importance of alternations between constricted and unconstricted settings in the Bai language. The distribution of laryngeal voice quality features in the infants' syllabic utterances also strongly suggested attunement to the target languages. By the end of the year, in months 10-12, the majority of syllabic utterances produced by English infants were unconstricted, while the majority of syllabic utterances produced by Bai infants were constricted or dynamic.

The infants' production of syllabic and non-syllabic utterances also displayed universal and language-specific properties. In line with previously reported studies of infant babbling (Koopmans-van Beinum & van der Stelt, 1986; Oller et al., 1998, 1999), all the infants in this study had begun to produce syllabic utterances by the seventh month of life. However, the development of syllabic utterances differed between the two language groups. Among the English infants, once syllabic utterances began to be produced in months 4-6, they steadily proliferated for the remainder of the year. Among Bai infants, syllabic utterances were produced early in infancy—towards the end of months 1-3 for two of the four Bai infants in the study—but did not increase significantly in the infants' vocal repertoires until months 10-12. This effect may reflect the influence of the target language. Based on previous research in infant speech perception (see, for example, Shi & Werker, 2001), infants attend to lexical words early in development, by six months of age. Given their role in creating lexical distinctions in Bai, it is likely that laryngeal voice quality variations were more salient to the Bai infants than to the English infants, who may have perceived the relatively greater importance of syllabic structures in creating the framework for lexical distinctions in their own ambient language. These differences may help to explain the different rate of development of syllabic utterances between the two language groups.

## 7.2 Adult Perceptions of Infant Vocalizations

The second component of the present study was an exploratory listening task, designed to examine Bai and English adults' perceptions of laryngeal voice quality in infant vocalizations. This study is unique in the literature, in that it focuses on laryngeal voice quality, spans a wide range of non-distress sounds produced by infants, and compares the perceptions of adults whose native languages contrast markedly in their use of laryngeal voice quality.

Bai and English adults listened to a range of constricted, dynamic, and unconstricted syllabic and non-syllabic sounds, and were asked to rate their importance in learning to speak their respective native languages. Bai adults did not strongly distinguish between utterances on the basis of laryngeal voice quality, most likely because lexical items in Bai may feature any combination of constricted, dynamic, or unconstricted voice quality settings. By contrast, English listeners showed a preference for unconstricted sounds, reflecting the dominance of unconstricted voice quality settings in their language. Compared to Bai listeners, English listeners also showed an aversion to constricted utterances, particularly in non-syllabic utterances, with ratings for dynamic utterances falling in between. The aversion to constriction may have partly reflected the use of contriction in English as a paralinguistic expression of negative emotion (Esling, 1994; Laver, 1980, 1994).

The ratings of adult Bai and English listeners reflected a consistent preference for syllabic over non-syllabic utterances, a pattern found in other perceptual studies (see, for example, Bloom & Lo, 1990, and Bloom et al., 1993). Among non-syllabic utterances, constricted syllabic utterances received the lowest mean ratings from listeners in both language groups. Both groups awarded higher ratings to dynamic and unconstricted non-syllabic utterances, with Bai listeners favouring dynamic utterances, and English listeners exhibiting a preference for unconstricted vocalizations. Throughout their ratings of non-syllabic sounds, whatever the laryngeal voice quality setting, Bai listeners preferred sounds that featured variations in pitch and laryngeal voice quality. English listeners tended to favour sounds produced in a moderate pitch range with smooth pitch contours and a single voice quality throughout—preferences that were particularly noticeable if the sound was produced with an unconstricted voice quality, as opposed to a dynamic or contricted quality.

While both groups of adult listeners preferred syllabic over non-syllabic utterances, Bai listeners did not make as sharp a distinction between syllabic and non-syllabic utterances as English listeners, consistently awarding non-syllabic utterances higher mean ratings than English listeners, except for some of the unconstricted non-syllabic sounds, where the ratings of the two language groups sometimes converged. The two groups differed most strikingly in their ratings of constricted non-syllabic utterances. While most English listeners gave these sounds the lowest ratings in the sample, many Bai listeners—especially male Bai listeners—rated these vocalizations quite highly. In the qualitative component of the study, many Bai listeners specifically cited the importance of laryngeal constriction in Bai, a factor that may account for the relatively high ratings of constricted non-syllabic sounds in this listening task.

Ratings of syllabic items for both language groups seemed to be made primarily on the basis of the infant’s ability to produce a rapidly articulated series of syllables, rather than on the basis of laryngeal voice quality. Among the English listeners, if an unconstricted syllabic utterance did not feature canonical syllable timing, it would receive a lower rating than a constricted syllabic utterance that did. However, unconstricted syllabic utterances with canonical timing received slightly higher ratings than constricted syllabic utterances with this feature. Similarly, unconstricted syllabic utterances with few, slowly produced syllables received higher mean ratings than constricted syllabic utterances with these characteristics. Thus, other things being (roughly) equal, English listeners showed a preference for unconstricted syllabic utterances, particularly those produced in a single voice quality throughout. By contrast, while, like the English listeners, Bai listeners favoured infants’ capacity to produce rapidly articulated syllables, the secondary factor influencing their ratings was the capacity to produce rapid alternations in pitch and/or voice quality, regardless of whether the utterance was constricted or unconstricted. Bai females favoured such alternations when they occurred from syllable to syllable, rather than gradually within an utterance, but Bai males were less selective, favouring variability in a more general sense.

Part of the purpose of the listening task was to determine whether adults could provide systematic evaluations of infant utterances, particularly non-syllabic utterances. The results of the perception test varied considerably, within and across language groups, and between items within utterance sub-categories. However, the variability proved to be largely systematic, suggesting that adults’ ratings were guided by genuine linguistic and phonetic preferences that could be further tested in larger, more controlled perceptual studies of infant utterances. Adults from both language groups tended to assign lower ratings to infant utterances that occur earlier in development, and higher ratings to sounds that occur later. For example, non-syllabic utterances received lower ratings than syllabic utterances. Within the larger category of non-syllabic utterances, adults gave the lowest ratings to constricted non-syllabic vocalizations, which infants are able to produce from birth, and they gave higher ratings to dynamic and unconstricted non-syllabic utterances, which occur later in development. Similarly, among syllabic utterances, those produced quickly with a large number of syllables received higher ratings than those produced more slowly with a smaller number of syllables; the latter occur earlier in development than the former. Finally, language-specific preferences in the ratings, such as the English preference for unconstricted utterances and the English aversion to constriction, were reflected in the patterns found in English infants' production. Similarly, the relatively high ratings given to non-syllabic utterances by the Bai adults were also reflected in the relatively high proportion of non-syllabic utterances produced by the Bai infants throughout most of the first year, compared to the English infants. Thus, adult perception reflects both universal and language-specific properties, as does infants' production.

## 7.3 Directions for Future Research

This dissertation reports on an exploratory study of the production and perception of laryngeal voice quality in the early vocalizations of Bai and English infants. This study is the first phonetically grounded, cross-linguistic study of infant production and adult perception that focuses on laryngeal voice quality. The results demonstrate the productivity of Esling's (2005) model of the vocal tract in highlighting universal and language-specific patterns in infants' phonetic development and in adults' perceptions of infant utterances. This study provides a strong foundation for future production- and perception-oriented research into laryngeal voice quality.

The findings of the production component of this study are based on an analysis of 2400 vocalizations of four Bai and four English infants. To confirm the findings of this study, it is important to extend the approach employed in this study to more Bai and English infants. Also, future research could focus on the production patterns of infants who are learning other register tone languages (for example, Yi, as documented by Esling & Edmondson, 2002), of infants that are learning languages where constriction is employed at the segmental rather than suprasegmental level (for instance, Arabic), and of infants who are learning other languages that do not employ laryngeal constriction contrastively at the segmental or suprasegmental level. This research would extend our understanding of universal and language-specific aspects of the development of laryngeal constriction in infancy. Also, the model piloted herein could be extended to the second year of life, when language-specific uses of laryngeal constriction may become stronger, particularly among infants who are learning register tone languages.

Finally, the findings of the perception study are based on 40 adult Bai and 40 adult English listeners' ratings of 36 sounds. Future studies could use a larger number of sounds within the sub-categories employed in this study, and stimuli within each of these sub-categories could be systematically selected to test for the relevance of some of the factors identified in this study as influencing adults' ratings. For example, within each of the subcategories, samples could be selected to reflect wide versus narrow pitch ranges, fast versus slow alternations in laryngeal voice quality, and so on. However, it is important to highlight the difficulty of conducting controlled perceptual studies of laryngeal voice quality using natural infant speech stimuli. Infant vocalizations range widely, and it is nearly impossible to find two samples that contrast only in laryngeal voice quality. Moreover, current technology does not allow us to manipulate infant speech stimuli to produce the full range of laryngeal voice quality settings explored in this study. Thus, other means of control may need to be employed in designing experimental research into the perception of laryngeal voice quality in infant vocalizations.

The qualitative component of the perception task highlighted a number of non-linguistic factors that likely affected adults' perceptions of laryngeal voice quality in this study. For example, adults cited their judgments about the age of the infants and the communicative purpose of the vocalizations as factors that may have affected their ratings (seeking attention, expressing emotion, and so on), as distinct from the question probed in the task, which was how important a given utterance was in learning the language. While it may not be possible to produce a set of infant speech stimuli that contrast purely in terms of laryngeal voice quality features, it may be fruitful to play the same infant utterances to different groups of adults within and across language groups, varying the type of judgments they are asked to make about the sounds. For example, if adults are told that all the sounds in a test are produced by five-month-old infants versus twelve-month-old infants, their ratings may differ considerably on the same sounds. Similarly, if one group of adults is asked to rate a group of sounds according to its perceived effectiveness in expressing emotion or seeking attention, that group may provide very different ratings of laryngeal voice quality than a group that is asked to rate the same sounds on the basis of their linguistic importance. By controlling conditions in this way, it may be possible to more precisely control some of the linguistic and non-linguistic factors that affect adults' perceptions of laryngeal constriction within and between linguistic and cultural groups.

Overall, the production and perception components of the present study, while explorative, provide a foundation for more systematic examination of the development of laryngeal constriction in infancy across languages, and for a more nuanced and controlled study of the factors that influence adults' perceptions of laryngeal constriction. The present study establishes such work as a productive direction for future investigation.

# Bibliography

Barr, R. G. (1990). The normal crying curve: What do we really know? *Developmental Medicine and Child Neurology*, *32*, 356-362.

Barr, R. G. (1998). Crying in the first year of Life: Good news in the midst of distress. *Child: Care, Health and Development, 24*, 425-439.

Benner, A., Grenon, I., & Esling, J. H. (2007). Infants' phonetic acquisition of voice quality parameters in the first year of life. In J. Trouvain & W.J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3* (pp. 2073-2076). Saarbrücken: Universität des Saarlandes.

Bettany, L. (2004). *Range exploration of phonation and pitch in the first six months of life*. Unpublished M.A. thesis, University of Victoria.

Bisping, R., Steingrueber, J., Oltmann, M., & Wenk, C. (1990). Adults' tolerance of cries: An experimental investigation of acoustic features. *Child Development*, *61*, 1218-1229.

Bloom, K., & Lo, E. (1990). Adult perceptions of vocalizing infants. *Infant Behavior and Development*, *13*, 209-219.

Bloom, K., D'Odorico, L., & Beaumont, S. (1993). Adult preferences for syllabic vocalizations: Generalizations to parity and native language. *Infant Behavior and Development*, *16*, 109-120.

Bloom, K., & Masataka, N. (1996). Japanese and Canadian impressions of vocalizing infants. *International Journal of Behavioral Development*, *19,* 89-99.

Boysson-Bardies, B. de. (1999). *How Language Comes to Children: From Birth to Two Years*. Cambridge MA: MIT Press.

Boysson-Bardies, B. de, Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language, 11*, 1-15.

Buder, E. H., Chorna, L. B., Oller, D. K., & Robinson, R. B. (2008). Vibratory regime classification of infant phonation. *Journal of Voice*, *22*(5), 553-564.

Buder, E. H., Oller, D. K., Magoon, J. C. (2003). Vocal intensity and phonatory regimes in the development of infant protophones. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 3*, 2014-2018.

Carlson, B. F., Esling, J. H., & Harris, J. G. (2004). A laryngoscopic phonetic study of Nlaka'pamux (Thompson) Salish glottal stop, glottalized resonants, and pharyngeals. In D. B. Gerdts & L. Matthewson (Eds.), *Studies in Salish linguistics in Honor of M. Dale Kinkade, Occasional Papers in Linguistics No. 17* (pp. 58-71). Missoula: University of Montana Press.

Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Clumeck, H. V. (1977). *Studies in the acquisition of Mandarin phonology*. Unpublished doctoral dissertation, University of California, Berkeley.

Clumeck, H. (1980). The acquisition of tone. In G. H. Yeni-Komshian, J. F. Kavanaugh, & C. A. Ferguson (Eds.), *Child Phonology: Vol. 1. Production* (pp. 257-275). New York: Academic Press.

Crelin, E. S. (1973). *Functional Anatomy of the Newborn*. New Haven: Yale University Press.

Davis, B., MacNeilage, P., Matyear, C., & Powell, J. (2000). Prosodic correlates of stress in babbling: An acoustical study. *Child Development*, *71*(5), 1258-1270.

Delack, J. B., & Fowlow, P. J. (1978). The ontogenesis of different vocalization: Development of prosodic contrastivity during the first year of life. In N. Waterson & D. Snow (Eds.), *The Development of Communication* (pp. 93-110). New York: Wiley.

Eckel, H. E., Koebke, J., Sittel, C., Sprinzl, G. M., Pototschnig, C., & Stennert, E. (1999). Morphology of the human larynx during the first five years of life studied on whole organ serial sections. *Annals of Otolaryngology, Rhinology and Laryngology*, *108*(3), 232-238.

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register, and stress: laryngoscopic case studies. *Phonology*, *23*(2), 157-191.

Edmondson, J. A., Esling, J. H., Harris, J. G., & Huang Tung-chiou. (2005). A laryngoscopic study of glottal and epiglottal/pharyngeal stop and continuant articulations in Amis - an Austronesian language of Taiwan. *Language and Linguistics*, *6*, 381-396.

Eimas, P. D. (1974). Auditory and linguistic processing of cues for place of articulation by infants.*Perception and Psychophysics*, *16*, 513-521.

Eimas, P. D., & Miller, J. L. (1980). Discrimination of the information for manner of articulation. *Infant Behavior and Development*, *3*, 367-375.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*, 303-306.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998a). “Babbling” and social context in infant monkeys: Parallels to human infants. *Trends in Cognitive Sciences*, *2*, 31-37.

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998b). Infant “babbling” in a nonhuman primate: Complex vocal sequences with repeated call types. *Behavior*, *135*, 643-664.

Emde, R. N., Gaensbauer, T. J., & Harmon, R. J. (1976). *Emotional expression in infancy: A biobehavioral study*. New York: International Universities Press.

Esling, J. H. (1994). Voice quality. In R. E. Asher, & J. M. Y. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (pp. 4950-4953). Oxford: Pergamon Press.

Esling, J. H. (1996). Pharyngeal consonants and the aryepiglottic sphincter. *Journal of the International Phonetic Association*, *26*, 65-88.

Esling, J. H. (1998). Laryngoscopic analysis of pharygeal articulations and layrnx-height voice quality settings. *Proceedings of the 5th International Conference on Spoken Language Processing, Vol. 2* (pp. 543-546). Sydney: ASSTA.

Esling, J. H. (1999a). The IPA categories “pharyngeal” and “epiglottal”: Laryngoscopic observations of pharyngeal articulations and larynx height. *Language & Speech*, *42*, 349-372.

Esling, J. H. (1999b). Voice quality setting of the pharynx. *Proceedings of the14th International Congress of Phonetic Sciences, 3*, 2449-2452.

Esling, J. H. (2000). Crosslinguistic aspects of voice quality. In R. D. Kent, & M. J. Ball (Eds.), *Voice Quality Measurement* (pp. 49-58). San Diego: Singular Publishing Group.

Esling, J. H. (2005). There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics*, *50*, 13-44.

Esling, J. H., Ball, M. J., & Dickson, B. C. (2000). The transcription of voice quality. In R. D. Kent, & M. J. Ball (Eds.), *Voice Quality Measurement* (pp. 49-58). San Diego: Singular Publishing Group.

Esling, J. H., Benner, A., & Bettany, L. (2004a). *Phonetic structure and acquisition of laryngeal and pharyngeal articulations: Text-analysis considerations*. Paper presented at the Canadian Symposium for Text Analysis, University of Victoria, November 18, 2003.

Esling, J. H., Benner, A., & Bettany, L. (2004b). *Phonetic articulatory control in the first year of life*. Paper presented at the 2004 Conference of the British Association of Academic Phoneticians, Cambridge University, March 24-26, 2004.

Esling, J. H., Benner, A., Bettany, L., & Zeroual, C. (2004c). Le contrôle articulatoire phonétique dans le prébabillage. In B. Bel & I. Marlien (Eds.), *Actes des XXVes journées d'étude sur la parole* (pp. 205-208). Fès: AFCP.

Esling, J. H., Benner, A., & Grenon, I. (2006). *Phonetic development in the first year of life: A comparison of English, Moroccan Arabic, and Bai infants*. Paper presented at the conference of the British Association of Academic Phoneticians, Edinburgh, April 2006.

Esling, J. H., & Edmondson, J. A. (2002). The laryngeal sphincter as an articulator: Tenseness, tongue root and phonation in Yi and Bai. In A. Braun & H. Masthoff (Eds.), *Phonetics and its Applications: Festschrift for Jens-Peter Koster on the Occasion of his 60th Birthday* (pp. 38-51). Stuttgart: Franz Steiner Verlag.

Esling, J. H., & Gao, M. (2003). Articulatory features of tones in whispered Chinese. *Proceedings of the 15$^{th}$ International Congress of Phonetic Sciences, Barcelona, 3,* 2629-2632.

Esling, J. H., & Harris, J. G. (2003) An expanded taxonomy of states of the glottis. *Proceedings of the 15$^{th}$ International Congress of Phonetic Sciences, Barcelona, 3*, 1049-1052.

Fitch, W.T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*, 1511-1522.

Fitch, W. T., Neubauer, J., & Herzel, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behavior, 63*, 407-418.

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London, 268*, 1669-1675.

Fleming, A. S., Corter, C., Stallings, J., & Steiner, M. (2002). Testosterone and prolactin are associated with emotional responses to infant cries in new fathers. *Hormones and Behavior, 42*, 399-413.

Foulkes, P., Docherty, G., & Watt, D. (2001). Tracking the emergence of structured variation. *University of of Pennsylvania Working Papers in Linguistics 6 – A Selection of Papers from NWAVE 27*.

Fried, M. P., Kelly, J. H., Strome, M. (1982). Comparison of the adult and infant larynx. *The Journal of Family Practice, 15*, 557-561.

Frodi, A. (1985). When empathy fails: Aversive infant crying and child abuse. In B. M. Lester & C. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 263-277). New York: Plenum Press.

Frodi, A. M., & Lamb, M. E. (1980). Child abusers' responses to infant smiles and cries. *Child Development*, *51*, 238–241.

Frodi, A., & Senchak, M. (1990). Verbal and behavioral responsiveness to the cries of atypical infants. *Child Development*, *61*, 76-84.

Furlow, F. B. (1997). Human neonatal cry quality as an honest signal of fitness. *Evolution and Human Behavior, 18*, 175-193.

Gao, M. (2002). *Tones in whispered Chinese: Articulatory features and perceptual cues*. Unpublished Master's thesis, University of Victoria.

Gao, M., & Esling, J. H. (2003). Articulatory features of tones in whispered Chinese. *Proceedings of the 15th International Congress of Phonetic Sciences,* Vol. 3, pp. 2629-2632. Barcelona: UAB.

Gergely, G., & Csibra, G., (2004). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Science*, *7*(7), 287-292.

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, *6*(3), 463-481.

Gilbert, H. R., & Robb, M. P. (1996). Vocal fundamental frequency characteristics of infant hunger cries : Birth to 12 months. *International Journal of Pediatric Otorhinolaryngology*, *34*, 237-243.

Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, *15*, 221-235.

Goberman, A. M., & Robb, M. P. (1999). Acoustic examination of preterm and full-term infant cries: The long-time average spectrum. *Journal of Speech, Language, and Hearing Research*, *42*, 850-861.

Goldfield, E. C. (2000). Exploration of vocal tract properties during serial production of vowels by full term and preterm infants. *Infant Behavior and Development, 23*, 421-439.

Goldman, H. I. (2001). Parental reports of “mama” sounds in infants: An exploratory study. *Journal of Child Language, 28*, 497-506.

Grauel, E. L., Höck, S., & Rothgänger, H. (1990). Jitter-index of the fundamental frequency of infant cry as a possible diagnostic tool to predict future developmental problems. Part 2: Clinical considerations. *Early Child Development and Care*, *65*, 23-29.

Grégoire, A. (1937). *L’apprentissage du langage: Les deux premières années*. Paris: Félix Alcan.

Grenon, I., Benner, A., & Esling, J. H. (2007). Language-specific phonetic production patterns in the first year of life. In J. Trouvain & W.J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences, vol. 3* (pp. 1561-1564). Saarbrücken: Universität des Saarlandes.

Hallé, P.A., Vihman, M.M. (1991). Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants. *Language & Speech*, *34*, 299-318.

Hauf, P., & Prinz, W. (2005). The understanding of own and others’ actions during infancy: “You-like-me” or “me-like-you”?. *Interaction Studies*, *6*(3), 429-435.

Hollien, H. (1974). On vocal registers. *Journal of Phonetics*, *2*. 125-143.

Hopkins, B. (2001). Development of crying: The origin and change problems. In R. G. Barr, I. St. James-Roberts, & M. R. Keefe (Eds.), *New Evidence on Unexplained Early Infant Crying: Its Origins, Nature and Management* (pp. 71-85). Calverton, NY: Johnson & Johnson Pediatric Institute.

Hopkins, B., & von Wulfften Palthe, T. (1987). The development of the crying state during early infancy. *Developmental Psychobiology, 20*, 165-175.

Hsu, Hui-Chin, Fogel, Alan & Cooper, Rebecca.  (2000). Infant vocal development in the first 6 months: Speech quality and melodic complexity. *Infant and Child Development*, *9*, 1-16.

Jakobson, R. (1941 /1968). *Child Language, Aphasia and Phonological Universals*. The Hague: Mouton.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants’ preference for the predominant stress patterns of English words. *Child Development, 64*, 675-687.

Jusczyk, P. W., & Kemler Nelson, D. G. (1996). Syntactic units, prosody, and psychological reality during infancy. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 389-408). Mahwah: Lawrence Erlbaum.

Keating, P. (1980). Patterns of fundamental frequency and vocal registers. In T. Murry & J. Murry (Eds.), *Infant Communication: Cry and Early Speech*. Houston: College-Hill.

Kent, R. D. (1981). Articulatory-acoustic perspectives on speech development in language behavior in infancy and early childhood. In R. E. Stark (Ed.), *Language Behavior in Infancy and Early Childhood* (pp. 105-126). New York: Elsevier North Holland.

Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, *72*, 353-365.

Kent, R. D., & Vorperian, H. K. (1995). Development of the craniofacial-oral-laryngeal anatomy: A review. *Journal of Medical Speech-Language Pathology*, *3*, 145-190.

Koopmans-van Beinum, F. J. & van der Stelt, J. M. (1986). Early stages in the development of speech movements. In B. Lindblom, & R. Zetterström (Eds.), *Precursors of Early Speech* (pp. 279-308). Basingstoke, Hampshire: MacMillan.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11850-11857.

Kuhl, P. K., Conboy, B., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the “critical period.” *Language Learning and Development*, *1*(3 & 4), 237-264.

Kuhl, P. K., & Meltzoff, A. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America*, *100*(4), 2425-2438.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, F13-F21.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World’s Languages*. Malden, MA: Blackwell.

Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, *4*, 185-199.

Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B. (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of Oral Biology*, *46*, 117-128.

Lieberman, P. (1985). The physiology of cry and speech in relation to linguistic behavior. In B. Lester & C. F. Z. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 37-52). New York: Plenum Press.

Lieberman, P., Harris, K. S., Wolff, P., & Russell, L. H. (1971). Newborn infant cry and nonhuman primate vocalization. *Journal of Speech and Hearing Research*, *14*, 718-727.

Locke, J. L. (1998). Social sound-making as a precursor to spoken language. In , J. R. Hurford, M. Studdert-Kennedy, and C. Knight (Eds.), *Approaches to the Evolution of Language: Social and Cognitive Bases* (pp. 190-201). Cambridge: Cambridge University Press.

Locke, J. L. (2006). Parental selection of vocal behavior: Crying, cooing, babbling, and the evolution of language. *Human Nature*, *17*(2), 155-168.

Lummaa, V., Vuorisalo, T., Barr, R. G., & Lehtonen, L. (1998). Why cry? Adaptive significance of intensive crying in human infants. *Evolution and Human Behavior*, *19*, 193-202.

MacNeilage, P. F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences*, *21*, 499-546.

MacNeilage, P. F., and Davis, B. (2001). The role of rhythmic cyclicities in infant action development. *Developmental Science*, *4*(1), 79-83.

MacNeilage, P. F., Davis, B., Kinney, A., & Matyear, C. (2000). The motor core of speech: A comparison of serial organization patterns in infants and languages. *Child Development*, *71*(1), 153-163.

Mattock, K., & Burnham, D. (2006). Chinese and English infants’ tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241-265.

Mattock, K., Molnar, M., Polka, L., & Burnham, D., (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, *106*, 1367-1381.

Ménard, L., & Boë, L.-J. (2004). L'émergence du système phonologique chez l'enfant; l'apport de la modélisation articulatoire. *Canadian Journal of Linguistics*, *49*(2), 155-174.

McCune, L., & Vihman, M. (2001). Early phonetic and lexical development: A productivity approach. *Journal of Speech, Language, and Hearing Research*, *44*, 670-684.

McCune, L., Vihman, M., Roug-Hellichius, L., Delery, D., & Gogate, L. (1996). Grunt communication in human infants (Homo sapiens). *Journal of Comparative Psychology*, *110*, 27-37.

Meadows, D., Elias, G., & Bain, J. (2000). Mothers' ability to identify infants' communicative acts consistently. *Journal of Child Language*, *27*, 393-406.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75-78.

Meltzoff, A. N., & Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, *54*, 702-709.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology*, *25*, 954-962.

Meltzoff, A. N., & Moore, M. K. (2002). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, *25*, 39-61.

Mergel, P., & Herzel, H. (1997). Modelling biphonation: The role of the vocal tract. *Speech Communication*, *22*, 141-154.

Moisik, S. R. (2008). *A three-dimensional model of the larynx and the laryngeal constrictor mechanism*. Unpublished Master's thesis, University of Victoria.

Moisik, S. R., & Esling, J. H. (2007). 3-D auditory-articulatory modeling of the laryngeal constrictor mechanism. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences*, vol. 1 (pp. 373-376). Saarbrücken: Universität des Saarlandes.

Möller, S., & Schönweiler, R. (1999). Analysis of infant cries for the early detection of hearing impairment. *Speech Communication*, *28*(3), 175-193.

Murray, A. D. (1985). Aversiveness is in the mind of the beholder: Perception of infant crying by adults. In B. M. Lester & C. F. Z. Boukydis (Eds.), *Infant Crying: Theoretical and Research Perspectives* (pp. 217-239). New York: Plenum Press.

Nathani, S., Ertmer, D., & Stark, R. (2006). Assessing vocal development in infants and toddlers. *Clinical Linguistics and Phonetics*, *20*(5), 351-369.

Nathani, S., & Oller, D. K. (2001). Beyond ba-ba and gu-gu: Challenges and strategies in coding infant vocalizations. *Behavior Research Methods, Instruments, & Computers, 33*(3), 321-330.

Nathani, S., Oller, D. K., & Cobo-Lewis, A. (2003). Final syllable lengthening (FSL) in infant vocalizations. *Journal of Child Language*, *30,* 3-25.

Oller, D. K. (1978). Infant vocalization and the development of speech. *Allied Health and Behavioral Sciences, 1*(4), 523-549.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. Yeni-Komshian, J., Kavanaugh, & C. Ferguson (Eds.), *Child Phonology* (pp.93-112). New York: Academic Press.

Oller, D. K. (2000). *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum.

Oller, D.K. (2004). Underpinnings for a theory of communicative evolution. In D. K. Oller & U. Griebel (Eds.), *Evolution of Communication Systems* (pp. 49-65). Cambridge, MA: The MIT Press.

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science*, *4*(1), 49-60.

Oller, D. K., Eilers, R. E., Neal, A. R., & Cobo-Lewis, A. B. (1998). Late onset canonical babbling: A possible early marker of abnormal development. *American Journal on Mental Retardation*, *103*, 249-263.

Oller, D. K., Eilers, R. E., Neal,, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication* Disorders, *32*, 223-245.

Ota, M. (2003). The development of lexical pitch accent systems: An autosegmental analysis. *Canadian Journal of Linguistics*, *48*(3/4), 357-383.

Papaeliou, C., Minadakis, G., & Cavouras, D. (2002). Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *Journal of Speech Language and Hearing Research, 45*, 311-317.

Papousek, M., & von Hofacker, N. (1998). Persistent crying in early infancy: A non-trivial condition of risk for the developing mother-infant relationship. *Child: Care, Health and Development*, *24*, 395-424.

Papousek, Wurmser, M. H., & von Hofacker, N. (2001). Clinical perspectives on unexplained early crying: Challenges and risks for infant mental health and parent-infant relationships. In R. G. Barr, I. St. James-Roberts, & M. R. Keefe (Eds.), *New Evidence on Unexplained Early Infant Crying: Its Origins, Nature and Management* (pp. 289-316). Calverton, NY: Johnson & Johnson Pediatric Institute.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421-435.

Robb, M. P. (2003). Bifurcations and chaos in the cries of full-term and preterm infants. *Folia Phoniatrica et Logopaedica*, *55*, 233-240.

Robb, M. P., Goberman, A. M., & Cacace, A. T. (1997). An acoustic template of newborn infant crying. *Folia Phoniatrica et Logopaedica*, *49*, 35-41.

Robb, M. P., & Saxman, J. H. (1988). Acoustic observations in young children's non-cry vocalizations. *Journal of the Acoustical Society of America*, *83*, 1876-1882.

Robb, M. P., Saxman, J. H., & Grant, A. A. (1989). Vocal fundamental frequency characteristics during the first two years of life. *Journal of the Acoustical Society of America*, *85*, 1708-1717.

Roug, L., Landberg, I., & Lundberg, L. J. (1989). Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *Journal of Child Language*, *16*, 19-40.

Rvachew, S., Creighton, D., Feldman, N., & Sauve, R. (2002). Acoustic-phonetic description of infant speech samples: Coding reliability and related methodological issues. *Acoustic Research Letters Online*, *3*, 24-28.

Sasaki, C. T., Levine, P. A., Laitman, J. T., & Crelin, E. S. (1977). Postnatal descent of the epiglottis in man. *Archives of Otolaryngology*, *103*, 169-171.

Scheiner, E., Hammerschmidt, K., Jürgens, U., & Zwirner, P. (2002). Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *Journal of Voice*, *16*(4), 509-529.

Serkhane, J. E., Schwartz, J. L., Boë, L. J., Davis, B. L., & Matyear, C. L. (2007). Infants' vocalizations analyzed with an articulatory model: A preliminary report. *Journal of Phonetics*, *35*, 321-340.

Sheppard, W. C., & Lane, H. L. (1968). Development of the prosodic features of infant vocalizing. *Journal of Speech & Hearing Research*, *11*, 94-108.

Shi, R., & Werker, J. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, *12*(1), 70-75.

Snow, David, & Balog, Heather. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, *112*, 1025-1058.

Stark, R. E. (1980). Prespeech segmental feature development. In G. Yeni-Komshian, J. Kavanaugh, & C. Ferguson (Eds.), *Child Phonology* (pp. 73-92). New York: Academic Press.

Stark, R. E., Bernstein, L. E. & Demarest, M. E. (1993). Vocal communication in the first 18 months of life. *Journal of Speech and Hearing Research*, 36, 548-558.

Stark, R., Rose, S., & McLagen, M. (1975). Features of infant sounds: The first eight weeks of life. *Journal of Child Language, 2*, 205-221.

Teshigawara, M. (2003). *Voices in Japanese animation : A phonetic study of vocal stereotypes of heroes and villains in Japanese culture*. Unpublished PhD dissertation, University of Victoria.

Thelen, E. (1991). Motor aspects of emergent speech: A dynamic perspective. In N.A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch, & M. Studdert- Kennedy (Eds.), *Biological and Behavioral Determinants of Language Development* (pp. 339-362). Hillsdale, NJ: Lawrence Erlbaum Associates.

Thodén, C. J., Järvenpää, A. L., & Michelsson, K. (1985). Sound spectrographic cry analysis of pain cry in prematures. In C. F. Z. Boukydis (Ed.), *Infant Crying: Theoretical and Research Perspectives* (pp. 105-117). New York: Plenum.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, *28*, 675-735.

Vihman, M. M., Nakai, S., DePaolis, R. A., & Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, *50*, 336–353.

Vorperian, H. K., Kent, R. D., Gentry, L. R., & Yandell, B. S. (1999). Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures: Preliminary results. *International Journal of Pediatric Otorhinolaryngology*, *49*, 197-206.

Watson, J. S. (1972). Smiling, cooing, and “the game.” *Merrill-Palmer Quarterly, 18*, 323-339.

Watson, J. S. (1979). Perception of contingency as a determinant of social responsiveness. In E. Thoman (Ed.), *Origins of the Infant’s Social Responsiveness* (pp. 33-64). Hillsdale, NJ: Erlbaum.

Watson, J. S. (1985). Continency perception in early social development. In T. Field & N. Fox, (Eds.), *The Origins of Social Responsiveness* (pp. 33-64). Hillsdale, NJ: Erlbaum.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development*, *1*(2), 197-234.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization in the first year of life. *Infant Behavior and Development*, *7*, 49-63.

Werker, J. F., & Tees, R. C. (1992). The organization and reorganization of human speech perception. *Annual Review of Neuroscience*, *15*, 377-402.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, *50*, 905-939.

Wermke, K., Mende, W., Manfredi, C., & Bruscaglioni, P. (2002). Developmental aspects of infant’s cry melody and formants. *Medical Engineering & Physics, 24*, 501-514.

Whalen, D. H., Levitt, A. G., & Goldstein, L. M. (2007). VOT in the babbling of French- and English-learning infants. *Journal of Phonetics*, *35*, 341–352.

Zeskind, P. S. (1987). Adult heart-rate responses to infant cry sounds. *British Journal of Developmental Psychology*, *5*, 73-79.

Zeskind, P. S., & Barr, R. G. (1997). Acoustic characteristics of naturally occurring cries of infants with “colic.” *Child Development*, *68*, 394-403.

Zeskind, P. S., & Shingler, E. A. (1991). Child abusers’ perceptual responses to newborn infant cries varying in pitch. *Infant Behavior and* Development, *14*, 335-347.

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language*, *27*, 3-42.